

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.62013 Seconds  
(without alignments)  
4310.252 Million cell updates/sec

Title: US-09-926-799-6  
Perfect score: 2896  
Sequence: 1 MKMASNDAPSDGAGLVLP.....VNPFYTLAPMGNGRRRIQ 548

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1953.5	67.5	539	1	COAT LORDV
2	1123.5	38.8	546	1	COAT SOUV3
3	284.5	9.8	2344	1	POLN_RHDV
4	281.5	9.7	702	1	COAT SMSV1
5	275.5	9.5	671	1	COAT FCVFG
6	269	9.3	703	1	COAT SMSV4
7	262	9.0	668	1	COAT FCVCG
8	261.5	9.0	668	1	COAT FCVFG
9	255.5	8.8	2208	1	POLN_MANCV
10	140	4.8	2193	1	POLG_CX16T
11	137.5	4.7	2194	1	POLG_HE701
12	137	4.7	2193	1	POLG_CX16G
13	128.5	4.4	2205	1	POLG_POL2W
14	126.5	4.4	2207	1	POLG_POL2L
15	125	4.3	2175	1	POLG_BOVEV
16	121.5	4.2	2183	1	POLG_CXB4E
17	119	4.1	2332	1	POLG_FMDVO
18	118.5	4.1	2183	1	POLG_CXB4J
19	116.5	4.0	901	1	POLG_ENMG3
20	116.5	4.0	2333	1	POLG_FMDV1
21	116	4.0	1197	1	CAMI_BRARE
22	116	4.0	2193	1	POLG_HE71M
23	114	3.9	2206	1	POLG_POL3L
24	112.5	3.9	932	1	CDG7_HUMAN
25	110.5	3.8	695	1	CICA_PAEPP
26	110.5	3.8	2201	1	POLG_CXA9
27	109.5	3.8	1324	1	SALI_HUMAN
28	109	3.8	2193	1	POLG_HE71B
29	109	3.8	2206	1	POLG_POL1M
30	108.5	3.7	834	1	POLG_ENMGO
31	108.5	3.7	1011	1	POLG_FMDVT
32	108.5	3.7	2063	1	NC06_HUMAN
33	108.5	3.7	2332	1	POLG_FMDVA

RESULT 1  
COAT LORDV  
ID COAT LORDV STANDARD; PRT; 539 AA.  
AC P54635;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Coat protein (Capsid protein).  
OS Lordsdale virus (Human enteric calicivirus).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=82658;  
RN (1)  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96005060; PubMed=7561776;  
RA Diggle, K.E., Hambden, P., Camilleri, O., Clarke, I.N.,  
Human enteric caliciviridae: the complete genome sequence and  
expression of viral life cycle particles from a genetic group II small  
round structured virus.  
RL J. Gen. Virol. 76:2349-2355 (1995).  
CC - SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC EMBL; X86557; CAA60255.1; -  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
KW Coat protein; Glycoprotein.  
FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 539 AA; 58775 MW; 585C3E7F2C5FD21 CRC64;  
Query Match 67.5%; Score 1953.5; DB 1; Length 539;  
Best Local Similarity 66.2%; Pred. No. 6.6e-136;  
Matches 365; Conservative 74; Mismatches 93; Indels 19; Gaps 4;  
QY 1 MKMASNDAPSDGAGLVLPINNEMALDPVAGAAIAAPLTGQNIIDPMTMNNFVQAP 60  
Db 1 MKMASNDAPSDGSAANLVPEVNEVMALEPVVGAIAAPVAGQNVDPWRNNFVQAP 60  
QY 61 GGEFTVSPRNPGEVLLNLNLELGPENPYLAHARMYNGYAGGFVQVVLGNAGFTAGKII 120  
Db 61 GGEFTVSPRNPAGEILWSAPLGPLNLYSLHSLRMVNGYAGGFVQVVLGNAGFTAGKVI 120  
QY 121 FAATPPNFPIDNLSAAQITMCPHVIVDVRQLEPVLIPDPVRRNFFHYNQSDSLRLTIA 180  
Db 121 FAAVPPNFPTEGLSPFSQVTMPFHIVDVRQLEPVLIPDPVRRNFFHYNQANDSTKLIA 180

ALIGNMENTS

Matches 250; Conservative 80; Mismatches 182; Indels 69; Gaps 17;

Qy 1 MKWASNDAAAPSNDGAG---LYPEINN-EMALDVPVAGAAIAAPLTGQQNIIDPWIMNPF 56  
Db 1 MMASKADAPASGASGAGQQLPEVNTADPLMEPVAGPTTAVATAGQNMIDPWVNNF 60

Qy 57 VQAPGGEFTVPRNSPQEVLLNLELGEINPYLAHLARMYNGYAGGFEVQVVLGNAFTA 116  
Db 61 VQSPQGEFTISPNTPDGILDFQLGPHLNPFLSHLSQMYNGVGNMVRILLAGNAFSA 120

Qy 117 GKIFAAIPNFPIDNLAAQITMCQPHVIVDVQLEPVLNMPDVRNNFHYNQSDSRL 176  
Db 121 GKIIIVCCVPGFTSSSLTIAQATLFPHVIADVRLTEPIEMPLEDVNRVLYHTNDNQFT-M 179

Qy 177 RLIMLYTPLR--ANNSGDDVFTVSCRVLTRPSPDFSENFVPPPTVESKTKPFTLPILTI 234  
Db 180 RLVCMLYTPLTRTGGGNSDSFVVAGKVLTPAPSDFSFLFVPTTIEQKTRAFVTPNIPL 239

Qy 235 SEMNSRFPPIESLHTSPTEINIVVQCGRVTLDGELMGTTLQLLPSQICAFRGVLTST 294  
Db 240 QTLNSRFPPLIQMILSPDASQVVFQNGKCLIDGQLLGTTPATSQLFVRGKINQGA 299

Qy 295 SRASQADATAPRLFNYYVHVQDLNLTGTPYDAEDIPGLTPTDERGK---VFGVASQ 350  
Db 300 -----RTLN-----LTVGDKPF-MAFDSAPVPGPDF-GKCDWMHRISKTP 339

Qy 351 RLSDTTTRAHAKVDTTAGFTPKLSLEISTDSDDDQ--NOPT-----KFT 396  
Db 340 NNTGSDPMRSVSVQTNVQGVFHLGSIQ-----FDEVFNHTGDYIGTIEWISQPS 392

Qy 397 PVGIGVDNEAEFQOWSLPDYSGQFTHNNMLAPAVAPNPPGQQLLFFRSQLPSSGGRS--N 454  
Db 393 PPG-----TDINLWEITPDYGSLSQAANLAPPVFPFGFGEALVYFVSAPFGPNRSAPN 446

Qy 455 GVLDCLVQEWVQHFYQESAAQTOVALRVYVNDTGKVLFEAKHLKGFMTANN--GD 512  
Db 447 DV-PCLLPQYITHFVSEQAFTMGDAALLHYVDPTNRNLGELFKLYGGVILTCVPNGVA 505

Qy 513 SPITVPPNGYFRFESVWNPFFYTLAPMTGN-----GRRRI 547  
Db 506 GPQQLPLNGVFLFVSVWSRFYQLKPVGTASTARGRLGVRI 546

RESULT 3  
POLN RHDV STANDARD; PRT; 2344 AA.  
ID POLN RHDV  
AC P27A10;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DE Non-structural polyprotein [Contains: RNA-directed RNA polymerase (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like protein); Coat protein].  
OS Rabbit hemorrhagic disease virus (RHDV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae; Lagovirus.  
OX NCBI\_TaxID=11976;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=9136157; PubMed=1840711;  
RA Meyers G., Wirblich C., Thiel H.-J.;  
RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide sequencing of a calicivirus genome.";  
RL Virology 184:664-676(1991).  
CC -I- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.  
CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA] (N).  
CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its

Qy 181 MLYTPLRANNSGDDVFTVSCRVLTRPSPDFSENFVPPPTVESKTKPFTLPILTISEMNS 240  
Db 181 MLYTPLRANNSGDDVFTVSCRVLTRPSPDFSENFVPPPTVESKTKPFTLPILTISEMNS 240

Qy 241 RFPVPIESLHTSPTEINIVVQCGRVTLDGELMGTTLQLLPSQICAFRGVLTSTSRASDQ 300  
Db 241 RFPVPIESLHTSPTEINIVVQCGRVTLDGELMGTTLQLLPSQICAFRGVLTSTSRASDQ 300

Qy 301 ADTATRLFNYYVHVQDLNLTGTPYDAEDIPGLTPTDERGK---VFGVASQ-RNLDSTTRA 359  
Db 296 -----SHDVTMLASQNSNYDTEEPAPLGTDPFVKIQGLLTQTTTRAGSIRA 346

Qy 360 HEAKVDTTAGFTPKLSLEISTD-SDDFDQDQPTKFTPVGI---GVDNEAEFQOWSLPD 415  
Db 347 HKATVSTGSHVHTFPKLSGVQFTDNTNDFQAGQNTKFTPVGIQDGDHQNPEQOWSLPN 406

Qy 416 YSQGFTHNNMLAPAVAPNPPGQQLLFFRSQLPSSGGRSNGVLDCLVPOEWVQHFYQESAP 475  
Db 407 YSGRTGHNHVLAPAVAPTFFGEQLLFFRSTMPGCSGYPNNMLDCLLPQEWVLFYQEAAP 466

Qy 476 AQTQVALRVYVNDTGKVLFEAKHLKGFMTANNGDSPTVPPNGYFRFESVWNPFFYTL 535  
Db 467 AQSDVALLRFVNDTGKVLFECKLHSGYITVAHTGFDYDLVLPNGYFRFESVWNPFFYTL 526

Qy 536 APMGTNGRRR 546  
Db 527 APMGTNGRRR 537

RESULT 2  
COAT\_SOUV3 STANDARD; PRT; 546 AA.  
ID COAT\_SOUV3  
AC Q04542;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE Coat protein (Capsid protein).  
OS Southampton virus (serotype 3).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae; Norovirus.  
OX NCBI\_TaxID=37129;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93142023; PubMed=8380940;  
RA Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;  
RT "Sequence and genome organization of a human small round-structured (Norwalk-like) virus.";  
RL Science 259:516-519(1993).  
CC -I- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; L07418; AAA92984.1; .  
DR PIR; B37491; B37491.  
DR InterPro; IPR004005; Calici\_coat.  
DR Pfam; PF00915; Calici\_coat.  
DR Rfam; RF00915; Calici\_coat; 1.  
KW Coat protein; Glycoprotein.  
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;  
Query Match 38.8%; Score 1123.5; DB 1; Length 546;  
Best Local Similarity 43.0%; Pred. No. 5.8e-75;





Db 508 QOTSRTLPASVG 519

RESULT 5

COAT FCVF9 STANDARD; PRT; 671 AA.

AC P27406;

DT 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 01-JUN-1994 (Rel. 29, Last annotation update)

DE Coat protein (Capsid protein).

GN C9P76.

OS Feline calicivirus (strain F9) (FCV).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Vesivirus.

OC NCBI\_TaxID=11981;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92410623; PubMed=1529544;

RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,

RA Turner P.C.;

RT "The complete nucleotide sequence of a feline calicivirus.";

RL Virology 190:443-448(1992).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=92339673; PubMed=1633955;

RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;

RA "Cloning and sequence determination of the feline calicivirus strain F9.";

RL Biochem. Soc. Trans. 20:26S-26S(1992).

CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M86379; AAA79327.1; -.

CC EMBL; Z11536; CRA77636.1; -.

CC PIR; B43382; VCWVF9.

CC InterPro; IPR004005; Calici\_coat.

CC InterPro; IPR008975; Viral\_cap\_coat.

CC Pfam; PF00915; Calici\_coat; 1.

CC Coat protein; Glycoprotein.

CC CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC SEQUENCE 671 AA; 73441 MW; 33BEE86D8370D5ES CRC64;

CC

Query Match 9.5%; Score 275.5; DB 1; Length 671;

Best Local Similarity 22.3%; Pred. No. 1.4e-12;

Matches 123; Conservative 70; Mismatches 174; Indels 185; Gaps 23;

QY 67 SPRNPGVLLMLELGPINPYLAHARMYNGAGFEVQVVLGNAGTAFTAGKIIPAAIPP 126

DB 179 STSETQKILFQSLGPLLNPYLEHLAKLYVAGSIEVRFSGVFGGKLAIVVPP 238

QY 127 NFPINLSAQITMCPHIVDQLEPNLMPDVRRNFFHVNQSDSLRLIAMLYTPL 186

DB 239 G--VDFVQSTMLQPHVLFDAQVEPVIFCLDURSTLYLHMSDTR--TSLVIMVNDL 295

QY 187 RANNSGDDVFTVSC--RVLTSPSPDFSNFLVPP-----TVES-----KT-- 224

DB 296 -INPYANDANSGCCIVTETKGPDKFHLKPKPGSMLTGHSIPSDLPKTSLLMIGNRY 354

QY 225 -----KPTTL-----PI-LTISEMSNSRFPPIESLHSP 253

DB 355 WSDITDFVIRPFVQANRHDFNQETAGWSTFRFRPISVTITEQNAKLGIGV-----A 408

QY 254 TENIVQCONG--RVTLDELGMGTTQLLPSSQICAFRGVLTTRSTSRASDAQDTATPRLFNY 311

DB 409 TDYIVFPIPDGMPDITIPGEL-----IPAGDYAITNGTGN----- 443

QY 312 YMHVQDLNLTGTPYDPAEDIPGLTDPFRGKVGFGVASQORNLDSTTRAHEAKVDTAGRF 371

DB 444 -----DIITATGYDTADIK--NNTNFRGMVYCGSLQRAWG-----DKKISNTAFIT 488

QY 372 TPKLGSLSTSDSD-----PDNQPTKFTPVGIGVDNEAEFQQWLSLPDYSQFTH 422

DB 489 TATL-----DGDNNKINPCNTIDQSKIVVFNQNHVGGKKAQTSDDTLALLGYTG----- 537

QY 423 NMLAPAVAFNFPGEQLL-----PFRSQLPSSGSRNGVLDCLVPQEWVH--FYQBS 473

DB 538 -----IGEQAIGSDRRVVRISTLPETGARGG-----NHPIFYKNS 573

QY 474 AP-----AQOVALRVYNP-----DTGKVLFEAKLHKLGPMT 506

DB 574 IKLGVIRSIDVFNSQLHTSRQLSLNHLVLLPPDSPFAVYRIIDNSGWSFIDGIDSGFSF 633

QY 507 IANNGDSPITVP 518

DB 634 VGVSGFGKLEFP 645

RESULT 6

COAT SMSV4 STANDARD; PRT; 703 AA.

AC P36285;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Coat protein (Capsid protein).

OS San Miguel sea lion virus (serotype 4) (SMSV 4).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Vesivirus.

OC NCBI\_TaxID=36407;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=92410750; PubMed=1529644;

RA Neill J.D.;

RT "Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus: identification of conserved and non-conserved amino acid sequences among calicivirus capsid proteins.";

RL Virus Res. 24:211-222(1992).

CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M87482; AAA16220.1; -.

CC PIR; C48562; C48562.

CC InterPro; IPR004005; Calici\_coat.

CC InterPro; IPR008975; Viral\_cap\_coat.

CC Pfam; PF00915; Calici\_coat; 1.

CC Coat protein; Glycoprotein.

CC CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).

CC SEQUENCE 703 AA; 77721 MW; C5DAD823B261073 CRC64;

CC

Query Match 9.3%; Score 269; DB 1; Length 703;

Best Local Similarity 32.9%; Pred. No. 4.4e-12;

Matches 71; Conservative 38; Mismatches 89; Indels 18; Gaps 7;

```
QY      8 AAPNDGAAGLVPE-----INNEAMALDPVAGAAIAAPLTGQOIIIDPWIMNMFVQAPGGE 63
DB      151 AESDGFSGAEIYTBESQGTVVQOQAPAPATLATATASTG-KSVGEQWMTFFSYHTSINW 209
QY      64 FTVSPRNSPCEVLLLELGPENIPYLALHARMYGVAGGFVQVVLGNAGFTAGKIIFAA 123
DB      210 STV---ESQCKLYSQALNPSINPYLDHAKLYSTWSGGIDVRFVTSVSGVFGGKLAALL 266
QY      124 IPPNF-PIDNLSAAQITMCPHVIVDVRQLEPNLPMPOVNRNFFHYNQGSLSRLRLIAML 182
DB      267 VPFGVEPIESVSMLOV---PHVLFDAQTEPVIPTIDIRKTLFRHMDTDT-TKLVINP 322
QY      183 YTPLRANSGDDVFTVSCRVLTRPSDPFSFNLVPP 218
DB      323 Y-----ENGVENKTKCSITVETRPSADFTFALLKPP 353
```

## RESULT 7

```
COAT_FCV66
ID COAT_FCV66 STANDARD; PRT; 668 AA.
AC P27404;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein (Capsid protein).
OS Feline calicivirus (strain CFI/68 FIV) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11979;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91374597; PubMed=1716692;
RA Neill J.D., Reardon I.M., Heinrichson R.L.;
RT "Nucleotide sequence and expression of the capsid protein gene of
   feline calicivirus."
RL J. Virol. 65:5440-5447 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Neill J.D.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; M32819; AAA42925.1; -.
DR EMBL; U13992; AAC13993.1; -.
DR PIR; A40507; VCNWFF.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 668 AA; 73550 MW; 9E2312108D4D441 CRC64;

Query Match          9.0%; Score 262; DB 1; Length 668;
Best Local Similarity 34.4%; Pred. No. 1.3e-11;
Matches 54; Conservative 32; Mismatches 59; Indels 12; Gaps 4;
```

```
QY      67 SPRNSPCEVLLLELGPENIPYLALHARMYGVAGGFVQVVLGNAGFTAGKIIFAAIYPP 126
DB      179 STSETQCKILFKQSLGLPLNPLYLTHLAKLYVAMSGVDVRFISGSGVFGGKLAIVVPP 238
```

```
QY      127 NFPIDNLSAAQITMCPHVIVDVRQLEPNLPMPOVNRNFFHYNQGSLSRLRLIAMLTYPL 186
DB      239 G-IDPVQSTSLQVPHVLFDAQTEPVIPTIDIRKTLFRHMDTDT-TSLVIMVYNDL 295
QY      187 -----RANSGDDVFTVSCRVLTRPSDPFSFNLVPP 218
DB      296 INPYANDSNSGCVITVE-----TKPGDPFKPHLLKPP 328
```

## RESULT 8

```
COAT_FCVF4
ID COAT_FCVF4 STANDARD; PRT; 668 AA.
AC P27405;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Coat protein (Capsid protein).
OS Feline calicivirus (strain Japanese F4) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11980;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91306470; PubMed=1853578;
RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,
   Miyamura K., Yamazaki S., Mikami T.;
RT "Sequence analysis of the 3'-end of feline calicivirus genome."
RL Virology 183:810-814 (1991).
RN [2]
RP -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
```

```
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
```

```
CC -----
DR EMBL; D90357; BAA14371.1; -.
DR PIR; B40481; VCNWFC.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 668 AA; 73589 MW; 85BDBC885804E503 CRC64;
```

```
Query Match          9.0%; Score 261.5; DB 1; Length 668;
Best Local Similarity 30.3%; Pred. No. 1.5e-11;
Matches 59; Conservative 41; Mismatches 78; Indels 17; Gaps 6;
```

```
QY      30 DPVAGAAIAAPLTGQOIIIDPW-IMNPFVQAPGGEFTVSPRNSPCEVLLLELGPENIY 88
DB      145 EPSAQMTAADVASGKSVDSWEAFPSFTS-----VNWSTSETQCKILFKQSLGLPLNRY 200
QY      89 LAHLARMYGVAGGFVQVVLGNAGFTAGKIIFAAIIPNFPIDNLSAAQITMCPHVIVD 148
DB      201 LEHLSKLYVAMSGSIEVRFSISGSGVFGGKLAIVVPPG--VDPVQSTSLQVPHVLFDA 258
QY      149 RQLEPNLPMPOVNRNFFHYNQGSLSRLRLIAMLTYPL-----RANSGDDVFTVSCRVL 203
DB      259 RQVEPIETIPDLRSTLYHVMSTDT-TSLVIMVYNDLINPYANDSNSGCVITVE----- 313
QY      204 TRPSDPFSFNLVPP 218
DB      314 TKPGDPFKPHLLKPP 328
```

Qy 2 KMSNDAAASNDGAGLVEINNEAMALDPVAGAAIAAPLTGQQNIIDPWIMNNFVQ--- 58

CC -|- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
 CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.  
 CC -|- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the  
 CC poliovirus polypeptide. In other picornavirus reactions Glu may be  
 CC substituted for Gln, and Ser or Thr for Gly.  
 CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC (RNA) (N).  
 CC -|- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -|- This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF177911; AAC55085.1; -.  
 CC HSP; P03300; IPOV.  
 CC InterPro; IPR003593; AAA\_ATPase  
 CC InterPro; IPR004004; Calici\_pol\_hel.  
 CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
 CC InterPro; IPR000199; Pept\_3C\_picorn.  
 CC InterPro; IPR000081; Peptidase\_C3.  
 CC InterPro; IPR003138; Pico\_P1A.  
 CC InterPro; IPR002527; Pico\_P2B.  
 CC InterPro; IPR001676; Rhv.  
 CC InterPro; IPR000605; RNA\_helicase.  
 CC InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 CC InterPro; IPR001205; RNA\_pol\_P3D.  
 CC InterPro; IPR007094; RNA\_pol\_P5vir.  
 CC InterPro; IPR008975; Viral\_cap\_coat.  
 CC Pfam; PF00548; Cys-protease\_3C; 1.  
 CC Pfam; PF02226; Pico\_P1A; 1.  
 CC Pfam; PF00947; Pico\_P2A; 1.  
 CC Pfam; PF01552; Pico\_P2B; 1.  
 CC Pfam; PF00073; rhv; 3.  
 CC Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 CC Pfam; PF00910; RNA\_helicase; 1.  
 CC PRINTS; PR00918; CALICIVIRUSNS.  
 CC ProDom; PD001125; Cys\_protease\_3C; 1.  
 CC ProDom; PD001306; Pico\_P2A; 1.  
 CC ProDom; PD001274; Pico\_P2B; 1.  
 CC SMART; SM00382; AAA; 1.  
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 CC Lipoprotein.  
 CC FT CHAIN 2 69 COAT PROTEIN VP4.  
 CC FT CHAIN 70 323 COAT PROTEIN VP2.  
 CC FT CHAIN 324 565 COAT PROTEIN VP3.  
 CC FT CHAIN 566 862 COAT PROTEIN VP1.  
 CC FT CHAIN 863 1012 CORE PROTEIN P2A.  
 CC FT CHAIN 1013 1111 CORE PROTEIN P2B.  
 CC FT CHAIN 1112 1440 CORE PROTEIN P2C.  
 CC FT CHAIN 1441 1526 CORE PROTEIN P3A.  
 CC FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VPg.  
 CC FT CHAIN 1549 1731 PICORNAIN 3C.  
 CC FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.  
 CC FT LIPID 2 2 N-myristoyl glycine (in host) (By  
 CC similarity).  
 CC FT ACT\_SITE 1695 1695 PROTEASE (POTENTIAL).  
 CC FT ACT\_SITE 1709 1709 PROTEASE (POTENTIAL).  
 CC SQ SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;  
 CC  
 CC Query Match 4.8%; Score 140; DB 1; Length 2193;  
 CC Best Local Similarity 21.9%; Pred.No. 0.063;  
 CC Matches 83; Conservative 41; Mismatches 133; Indels 122; Gaps 18;  
 CC  
 CC 57 VQAPGGFTVSPRNSPGFVLLNLELGPINPYLAHLARMYNGYAGGFVQVVLGNAFTA 116

Db 398 VOSKTGELCAAFRADPGRD-----GPMQSTILGQLCRYTQWSGSLVTFMFAGSFMT 451  
 Qy 117 GKIIIPAAIPP--NFPIDNLAAQITMCQPHVIVDVVQLEPVLNPMQDVRNPNPHYQGS 174  
 Db 452 GKMLIAYTPPGCNVPADRTAMLGT--HVIWDFGLQSSVTLVVPWISNT--HYR----- 501  
 Qy 175 RLRLIAMLYPLRANSGDVFTVSCRVLTRSPDFSNFLVP-----PT-----VESKTK 225  
 Db 502 -----AHARAGVFDYTTGIIITW-----YQTNVVPFGAPTAYIVALAAQD 545  
 Qy 226 PFTLPILLITSEMSNSRFPVPIE---SLHTSPTEIVVQCGRVTLDGELMGTTLQSPQ 282  
 Db 546 NPTMKLCKDTE-----DIEQTANIQGDPIADMIDQTVNNQV---NRSLTALQVLP-- 592  
 Qy 283 ICAPRGVLTSTRASDAQDQATATPRLFNVYHVQLDNLNGTYPDPAEDIPGLGT---PD 339  
 Db 593 -----TAADTEASSHR-----LGTGVVPA 611  
 Qy 340 FRGKVFGV---ASQNLDSLT--TRAHAKVDTTAGRTPKLGLSLISTSDDDFDQNP 394  
 Db 612 LQAETGASSNASDKNLIETRCVLNHHSTQETAGNFFSRAGLVSIIT-----MPT- 662  
 Qy 395 FFPVIGIGVDNEAEFOQWSL 413  
 Db 663 -----TGQTNTDGYVNWDI 676  
 RESULT 11  
 POLG\_HE701  
 ID POLG\_HE701 STANDARD; PRT; 2194 AA.  
 AC P32537;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polypeptide [Contains: Coat protein VP4 (P1A); Coat protein VP2  
 DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein  
 DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
 DE linked protein VPg (P3B); Picornain 3C (SC 3.4.22.28) (Protease 3C)  
 DE (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
 OS Human enterovirus 70 (strain J670/71) (EV 70).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 OC NCBI TaxID=31915;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91037960; PubMed=2172447;  
 RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D.,  
 RA Minor P.D., Almond J.W.;  
 RT "The complete nucleotide sequence of enterovirus type 70;  
 RT relationships with other members of the picornaviridae.";  
 RL J. Gen. Virol. 71:2291-2299(1990).  
 CC -|- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
 CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.  
 CC -|- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the  
 CC poliovirus polypeptide. In other picornavirus reactions Glu may be  
 CC substituted for Gln, and Ser or Thr for Gly.  
 CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC (RNA) (N).  
 CC -|- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AF177911; AAC55085.1; -.  
 CC HSP; P03300; IPOV.  
 CC InterPro; IPR003593; AAA\_ATPase  
 CC InterPro; IPR004004; Calici\_pol\_hel.  
 CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
 CC InterPro; IPR000199; Pept\_3C\_picorn.  
 CC InterPro; IPR000081; Peptidase\_C3.  
 CC InterPro; IPR003138; Pico\_P1A.  
 CC InterPro; IPR002527; Pico\_P2B.  
 CC InterPro; IPR001676; Rhv.  
 CC InterPro; IPR000605; RNA\_helicase.  
 CC InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 CC InterPro; IPR001205; RNA\_pol\_P3D.  
 CC InterPro; IPR007094; RNA\_pol\_P5vir.  
 CC InterPro; IPR008975; Viral\_cap\_coat.  
 CC Pfam; PF00548; Cys-protease\_3C; 1.  
 CC Pfam; PF02226; Pico\_P1A; 1.  
 CC Pfam; PF00947; Pico\_P2A; 1.  
 CC Pfam; PF01552; Pico\_P2B; 1.  
 CC Pfam; PF00073; rhv; 3.  
 CC Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 CC Pfam; PF00910; RNA\_helicase; 1.  
 CC PRINTS; PR00918; CALICIVIRUSNS.  
 CC ProDom; PD001125; Cys\_protease\_3C; 1.  
 CC ProDom; PD001306; Pico\_P2A; 1.  
 CC ProDom; PD001274; Pico\_P2B; 1.  
 CC SMART; SM00382; AAA; 1.  
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 CC Lipoprotein.  
 CC FT CHAIN 2 69 COAT PROTEIN VP4.  
 CC FT CHAIN 70 323 COAT PROTEIN VP2.  
 CC FT CHAIN 324 565 COAT PROTEIN VP3.  
 CC FT CHAIN 566 862 COAT PROTEIN VP1.  
 CC FT CHAIN 863 1012 CORE PROTEIN P2A.  
 CC FT CHAIN 1013 1111 CORE PROTEIN P2B.  
 CC FT CHAIN 1112 1440 CORE PROTEIN P2C.  
 CC FT CHAIN 1441 1526 CORE PROTEIN P3A.  
 CC FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VPg.  
 CC FT CHAIN 1549 1731 PICORNAIN 3C.  
 CC FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.  
 CC FT LIPID 2 2 N-myristoyl glycine (in host) (By  
 CC similarity).  
 CC FT ACT\_SITE 1695 1695 PROTEASE (POTENTIAL).  
 CC FT ACT\_SITE 1709 1709 PROTEASE (POTENTIAL).  
 CC SQ SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;  
 CC  
 CC Query Match 4.8%; Score 140; DB 1; Length 2193;  
 CC Best Local Similarity 21.9%; Pred.No. 0.063;  
 CC Matches 83; Conservative 41; Mismatches 133; Indels 122; Gaps 18;  
 CC  
 CC 57 VQAPGGFTVSPRNSPGFVLLNLELGPINPYLAHLARMYNGYAGGFVQVVLGNAFTA 116



DR ProDom; PD001125; Cys\_protease\_3C; 1.  
 DR ProDom; PD001306; Pico\_P2A; 1.  
 DR ProDom; PD001274; Pico\_P2B; 1.  
 DR SMART; SM00382; AAA; 1.  
 KW Polypeptide; Coat protein; Core protein; Transferase;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 KW Lipoprotein.  
 FT CHAIN 2 69 COAT PROTEIN VP4.  
 FT CHAIN 70 323 COAT PROTEIN VP2.  
 FT CHAIN 324 565 COAT PROTEIN VP3.  
 FT CHAIN 566 862 COAT PROTEIN VP1.  
 FT CHAIN 863 1012 CORE PROTEIN P2A.  
 FT CHAIN 1013 1111 CORE PROTEIN P2B.  
 FT CHAIN 1112 1440 CORE PROTEIN P2C.  
 FT CHAIN 1441 1526 CORE PROTEIN P3A.  
 FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VP3.  
 FT CHAIN 1549 1731 PICORNAIN 3C.  
 FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.  
 FT LIPID 2 2 N-myristoyl glycine (in host) (By similarity).  
 FT ACT SITE 1695 1695 PROTEASE (POTENTIAL).  
 FT ACT SITE 1709 1709 PROTEASE (POTENTIAL).  
 SQ SEQUENCE 2193 AA; 243209 MW; 04B38CE572A76E38 CRC64;

Query Match 4.7%; Score 137; DB 1; Length 2193;  
 Best Local Similarity 20.6%; Pred. No. 0.1;  
 Matches 92; Conservative 48; Mismatches 167; Indels 140; Gaps 20;

QY 57 VQAPGGEFTVSRNPGVLLNLEGLPEINPVLHARMYNGVAGFEVQVLAGNAFTA 116  
 DB 398 VQSKTGELCAVFRADPG-----RNGPWQSTLGLCRYTQWGSLEVTFFAGSEFMAT 451  
 QY 117 GKIIIFAAIAPP--NFPIIDLSAAQITMCPHVIIVDRQLEPVNLPMPDVRNPNFFHYQGSDS 174  
 DB 452 GKMLIATYPPGGVPAADRLTAMLTG---HVINDFGLQSSVTLVPIWISNT--HYR----- 501  
 QY 175 RURLIAMLYTPRANSGDDVTVCVRVTRSPDFSNFLVPPVTSKTFEFTLPILTI 234  
 DB 502 -----AHAKDGYDYTTTGTITW-----YQTNVYVP--IGAPTAYIVALLAA 543  
 QY 235 SEMNSRFPVPVIESLHSTP-----ENIVQCGNRTLDGELMGTTLQLLSQICAFRG 288  
 DB 544 QNFTWKLCKDTEDIEQSANIOGDGIADMDQAVTSRV---GRALTSLQVEP----- 592  
 QY 289 VLTRTSRASQDQATATPRLFNHYVHVDNLNLTGTPYDAEDIPGLPTG---PDRFGKVF 345  
 DB 593 ---TAANTNASEH-----LGTGLVPALQAAET 617  
 QY 346 GVAS---QRNLDT--TRAHEAKVDTAGRTFKGLSLEISTSDDDFQDQNTKFTPVGI 400  
 DB 618 GASSNAQDENLITRCVLNHHSTQETTIGNFFSRAGLVSIIT-----MPT-----T 663  
 QY 401 GVDNEAFQWSLPDYSGQ-----FTHNNLAPAVAPNPPGEGQLLFFRSQLPSSGG 451  
 DB 664 GTQNTDGYNNWDI--DLMGVAQMRKCELFY-----MRFOAEFTFVA 705  
 QY 452 RSGVGLDCLVPOEWQHFQESAPACT 478  
 DB 706 KPNGE---LVPQLLQYMYVPPGAPKET 729

## RESULT 13

POLG\_POL2W  
 ID POLG\_POL2W STANDARD; PRT; 2205 AA.  
 AC P23069;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polypeptide [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Picornain 2A (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].

OS Poliovirus type 2 (strain W-2).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 OX NCBI\_TaxID=12085;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90155230; PubMed=2154539;  
 RA Pevear D.C., Oh C.K., Cunningham L.L., Calenoff M., Jubelt B.;  
 RT "Localization of genomic regions specific for the attenuated, mouse-  
 adapted poliovirus type 2 strain W-2";  
 RL J. Gen. Virol. 71:43-52(1990).  
 CC FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
 CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
 CC CATALYTIC ACTIVITY: Selective cleavage of Gln-Ile bond in the  
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
 CC substituted for Gln, and Ser or Thr for Gly.  
 CC CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC (RNA) (N).  
 CC SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 EMBL; D00625; BAA00516.1; ALT\_SEQ.  
 PIR; A34032; GNNY2W.  
 HSP; P03299; LPOV.  
 MEROPS; C03.020; --  
 InterPro; IPR003593; AAA ATPase.  
 InterPro; IPR004004; Calici\_pol\_hel.  
 InterPro; IPR009003; Cys Ser trypsin.  
 InterPro; IPR001199; Pept\_3C\_picorn.  
 InterPro; IPR000081; Peptidase\_C3.  
 InterPro; IPR003138; Pico\_P1A.  
 InterPro; IPR002527; Pico\_P2B.  
 InterPro; IPR001676; Rhv.  
 InterPro; IPR006055; RNA helicase.  
 InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 InterPro; IPR001205; RNA\_pol\_P3D.  
 InterPro; IPR007094; RNA\_pol\_Psvir.  
 InterPro; IPR008975; Viral\_cap\_coat.  
 Pfam; PF00548; Cys-protease-3C; 1.  
 Pfam; PF02226; Pico\_P1A; 1.  
 Pfam; PF00947; Pico\_P2A; 1.  
 Pfam; PF01552; Pico\_P2B; 1.  
 Pfam; PF000073; rhv; 3.  
 Pfam; PF00680; RNA dep RNA\_pol; 1.  
 Pfam; PF00910; RNA helicase; 1.  
 PRINTS; PR00918; CALICIVIRUSN.  
 ProDom; PD001125; Cys\_protease\_3C; 1.  
 ProDom; PD001306; Pico\_P2A; 1.  
 ProDom; PD001274; Pico\_P2B; 1.  
 SMART; SM00382; AAA; 1.  
 KW Polypeptide; Coat protein; Core protein; Transferase;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 KW Lipoprotein.  
 FT CHAIN 2 69 COAT PROTEIN VP4.  
 FT CHAIN 70 340 COAT PROTEIN VP2.  
 FT CHAIN 341 578 COAT PROTEIN VP3.  
 FT CHAIN 579 879 COAT PROTEIN VP1.  
 FT CHAIN 880 1028 PROTEASE 2A.  
 FT CHAIN 1029 1125 CORE PROTEIN 2B.  
 FT CHAIN 1126 1454 CORE PROTEIN 2C.  
 FT CHAIN 1455 1541 CORE PROTEIN 3A.







FT	STRAND	83	87	FT	STRAND	542	546
FT	TURN	88	89	FT	TURN	550	551
FT	STRAND	90	94	FT	STRAND	574	574
FT	STRAND	101	102	FT	STRAND	578	578
FT	HELI	103	105	FT	STRAND	582	582
FT	TURN	113	115	FT	TURN	595	596
FT	STRAND	123	123	FT	STRAND	597	598
FT	STRAND	126	128	FT	HELI	600	602
FT	TURN	129	129	FT	HELI	610	613
FT	STRAND	133	134	FT	HELI	619	619
FT	STRAND	138	139	FT	STRAND	624	624
FT	TURN	142	143	FT	STRAND	626	628
FT	STRAND	147	151	FT	HELI	629	629
FT	TURN	152	152	FT	STRAND	630	634
FT	HELI	153	155	FT	HELI	638	641
FT	TURN	156	157	FT	STRAND	644	644
FT	HELI	159	167	FT	TURN	647	649
FT	STRAND	168	180	FT	STRAND	651	655
FT	TURN	185	186	FT	HELI	662	668
FT	STRAND	187	197	FT	TURN	669	670
FT	TURN	198	198	FT	STRAND	671	685
FT	HELI	209	212	FT	STRAND	687	687
FT	HELI	215	217	FT	TURN	689	690
FT	STRAND	219	220	FT	STRAND	693	693
FT	HELI	224	226	FT	STRAND	699	705
FT	TURN	227	230	FT	TURN	707	708
FT	HELI	234	238	FT	HELI	718	721
FT	STRAND	241	245	FT	STRAND	727	731
FT	TURN	246	248	FT	TURN	732	733
FT	STRAND	251	256				
FT	STRAND	265	265				
FT	TURN	267	269				
FT	STRAND	270	270				
FT	STRAND	273	285				
FT	STRAND	294	309				
FT	TURN	325	328				
FT	TURN	332	333				
FT	STRAND	340	340				
FT	TURN	343	344				
FT	STRAND	356	357				
FT	STRAND	359	359				
FT	TURN	360	360				
FT	HELI	361	364				
FT	TURN	365	365				
FT	STRAND	368	369				
FT	STRAND	375	375				
FT	TURN	376	377				
FT	TURN	380	381				
FT	HELI	382	384				
FT	STRAND	386	389				
FT	TURN	394	395				
FT	STRAND	397	402				
FT	TURN	407	408				
FT	HELI	410	413				
FT	HELI	415	420				
FT	TURN	421	422				
FT	STRAND	423	427				
FT	STRAND	430	436				
FT	TURN	440	441				
FT	STRAND	443	443				
FT	STRAND	445	451				
FT	HELI	461	464				
FT	TURN	465	466				
FT	STRAND	468	473				
FT	STRAND	479	484				
FT	STRAND	493	494				
FT	TURN	499	502				
FT	HELI	503	506				
FT	STRAND	510	515				
FT	STRAND	520	520				
FT	TURN	523	524				
FT	STRAND	528	537				
FT	TURN	539	540				

Query Match 4.3%; Score 125; DB 1; Length 2175;  
Best Local Similarity 19.8%; Pred. No. 0.79; Indels 64; Gaps 10;  
Matches 66; Conservative 47; Mismatches 156;  
QY 3 MASNDAPSNDGAAGLVPEINNEAMALDPVAGAAIAAAPTQQQNIIDPWNNFVQAPGG 62  
Db 313 IASNOGLTKPGPG-----SYQFWTDEDPCILPDFQPTLFIFIPGKVNLLLEIAQV 366  
QY 63 EFTVSPRNSPGE-----VLINLEL-----GPEINPYLAHLARMYNGY 99  
Db 367 ESILEANNREGVEGVYVIVSVQDALDAQIYALRLELGGSGPLSSLLGTAKHYTQM 426  
QY 100 AGGFEVQVVLAGNAFTAGKIIFAAIPNFPIDNLSAAQITMCPHVIVDVRLQEPVNLPM 159  
Db 427 SGSVEITCMFTGTFTMTGKVLAYTPGGMFR-NREAMLGTHVWVDFGLOSSTLVIP 485  
QY 160 DVNNFFHYNQSDSRRLRLIAMLYTPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPT 219  
Db 486 WISASHFR-----GVSNDDVLNYQYAAAGHVITWTQNMVIPPG 524  
QY 220 VESKTKPTLPLITISEMSNSRFPVPIESLHTSPTENIVQCONGRV---TLDGELMGTT 276  
Db 525 F-----PNTAGIIMMI-AAQPNFSFRIQKREDMTQTALQNDPGKMLKDAIDKOVAGA- 577  
QY 277 QLLPSQICAFRCGLTRSTS--RASDAQDTATPR 307  
Db 578 -LVAGTTTSTHSVATDSTPALQQAETGAISTAR 609

Search completed: June 1, 2004, 13:47:18  
Job time : 8.62013 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 31.9054 Seconds  
(without alignments)  
5419.275 Million cell updates/sec

Title: US-09-926-799-6  
Perfect score: 2896  
Sequence: 1 MKMASNDAPSDGAGLVLP.....VNPFYTLAPMTGNGRRRIQ 548

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp Vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2855	98.6	548	12 Q8BC87	Q8bc87 human calic
2	2851	98.4	548	12 Q8PYA7	Q8pya7 human calic
3	2850	98.4	548	12 Q916E5	Q916e5 human calic
4	2836	97.9	548	12 Q8BC90	Q8bc90 human calic
5	2831	97.8	548	12 Q915D2	Q915d2 human calic
6	2822	97.4	548	12 Q66296	Q66296 calicivirus
7	2819	97.3	548	12 Q915D3	Q915d3 human calic
8	2815	97.2	548	12 Q8V0P5	Q8v0p5 human calic
9	2812	97.1	548	12 Q918A1	Q918a1 norwalk-lik
10	2810	97.0	548	12 Q8V0P4	Q8v0p4 human calic
11	2802	96.8	548	12 Q8V768	Q8v768 norwalk vir
12	2799	96.7	548	12 Q88291	Q88291 small round
13	2794	96.5	548	12 Q83880	Q83880 norwalk vir
14	2786	96.2	548	12 Q91725	Q91725 norwalk-lik
15	2781	96.0	548	12 Q91722	Q91722 norwalk-lik
16	2780	96.0	548	12 Q91V40	Q91v40 human calic

17	2777	95.9	548	12 Q8V771	Q8v771 norwalk vir
18	2776	95.9	548	12 Q91728	Q91728 norwalk-lik
19	2772	95.7	548	12 Q68291	Q68291 human calic
20	2770	95.6	548	12 Q918A4	Q918a4 norwalk-lik
21	2768	95.6	534	12 Q9QAX4	Q9qax4 norwalk-lik
22	2742	94.7	548	12 Q91V46	Q91v46 human calic
23	2727	94.2	548	12 Q80FK9	Q80fk9 human calic
24	2720	93.9	548	12 Q80FK8	Q80fk8 human calic
25	2720	93.9	548	12 Q80FK5	Q80fk5 human calic
26	2718	93.9	548	12 Q80FK6	Q80fk6 human calic
27	2715	93.8	548	12 Q80FL1	Q80fl1 human calic
28	2707	93.5	548	12 Q80FK3	Q80fk3 human calic
29	2704	93.4	548	12 Q80FK7	Q80fk7 human calic
30	2701	93.3	548	12 Q80FK4	Q80fk4 human calic
31	2695	93.1	548	12 Q80FK2	Q80fk2 human calic
32	2688	92.8	548	12 Q80FL2	Q80fl2 human calic
33	2688	92.8	548	12 Q80FK1	Q80fk1 human calic
34	2676	92.4	548	12 Q80FL0	Q80fl0 human calic
35	2173.5	75.1	547	12 Q918A7	Q918a7 norwalk-lik
36	2171.5	75.0	547	12 Q8V775	Q8v775 norwalk vir
37	2161.5	74.6	547	12 Q8V773	Q8v773 norwalk vir
38	2148.5	74.2	535	12 Q917X4	Q917x4 norwalk-lik
39	2147.5	74.2	535	12 Q916E6	Q916e6 human calic
40	2147.5	74.2	535	12 Q916E4	Q916e4 human calic
41	2145.5	74.1	535	12 Q917Y3	Q917y3 norwalk-lik
42	2140.5	73.9	535	12 Q68104	Q68104 hawaii cali
43	2140.5	73.9	535	12 Q917Y9	Q917y9 norwalk-lik
44	2139	73.9	550	12 Q918B3	Q918b3 norwalk-lik
45	2137.5	73.8	535	12 Q917Y0	Q917y0 norwalk-lik

#### ALIGNMENTS

#### RESULT 1

Q8BC87 PRELIMINARY; PRT; 548 AA.  
 AC Q8BC87;  
 DT 01-MAR-2003 (Tremblrel. 23, Created)  
 DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
 DE Capsid protein.  
 OS Human calicivirus NLV/Oberhausen 455/01/DE.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OC NCBI\_TaxID=173921;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/Oberhausen 455/01/DE;  
 RA Kuenkel U., Schreier E.;  
 RT "Characterization of recombinant human calicivirus circulating in Germany."  
 RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF539440; AAN16373.1;  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 548 AA; 59985 MW; 9797ED4783790887 CRC64;

Query Match 98.6%; Score 2855; DB 12; Length 548;  
 Best Local Similarity 97.8%; Pred. No. 9.3e-218;  
 Matches 536; Conservative 9; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKMASNDAPSDGAGLVLPPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNFFVOAP 60  
 DB 1 MKMASNDATPNDGAGLVLPPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNFFVOAP 60.  
 QY 61 GGEFTVSPRSPGEVLLNLELGPENPYLAHLARMYNGYAGFEVQVVLGNAGNFTAGKII 120  
 DB 61 GGEFTVSPRSPGEVLLNLELGPENPYLAHLARMYNGYAGFEVQVVLGNAGNFTAGKII 120  
 QY 121 FAAPPNPPIINLSAQITMCPHVIVDVRQLEPNLPMDFVNNFFHNNQGSDSLRLIA 180  
 |||||

Db 121 FAAPNPFIDNLSAAQITMCPHVIVDVRLQEPVNLPMFVDRNNFFHYNQSDSRLRLIA 180  
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPPTVESKTKPFTLPILITISEMSNS 240  
Db 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPPTVESKTKPFTLPILITISEMSNS 240  
QY 241 RFPVPISLHTSPTENIVVQCONGRVTLDELGMGTQQLPSQICAFRGVLTRSTRASDQ 300  
Db 241 RFPVPISLHTSPTENIVVQCONGRVTLDELGMGTQQLPSQICAFRGVLTRSTRASDQ 300  
QY 301 ADTATPRLFNYYWHVQDNLNGTTPYDPAEDIPGLGTPDFRGKVFQVASQORNLDSSTTRAH 360  
Db 301 ADTATPRLFNYYWHVQDNLNGTTPYDPAEDIPGLGTPDFRGKVFQVASQORNLDSSTTRAH 360  
QY 361 EAKVDTTACGFTPKLGSLEISTDSDDFQDQNTKFTTPVGIGVDNEAEFQWMLSDPYSGOF 420  
Db 361 EAKVDTTACGFTPKLGSLEISTDSDDFQDQNTKFTTPVGIGVDNEAEFQWMLSDPYSGOF 420  
QY 421 THNNMLAPAVAPNFPGEQLLFFRSQLPSSGGRSNGVLDCLVQEWVQHFYQESAPAQTOV 480  
Db 421 THNNMLAPAVAPNFPGEQLLFFRSQLPSSGGRSNGVLDCLVQEWVQHFYQESAPAQTOV 480  
QY 481 ALVRYNPDGTGKVLFEAKLHKLGFMTIANNQSDSPITVPPNGYFRFESWNPFTYTLAPMGT 540  
Db 481 ALVRYNPDGTGKVLFEAKLHKLGFMTIANNQSDSPITVPPNGYFRFESWNPFTYTLAPMGT 540  
QY 541 GNGRRRIQ 548  
Db 541 GNGRRRIQ 548

RESULT 2  
Q9PYA7 PRELIMINARY; PRT; 548 AA.  
ID Q9PYA7  
AC Q9PYA7  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus strain Arg320.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OC NCBI\_TaxID=106586;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Arg320;  
RX MEDLINE=20129490; PubMed=10664391;  
RA Jiang X., Espul C., Zhong W.M., Cuello H., Matson D.O.;  
RT "Characterization of a novel human calicivirus that may be a naturally occurring recombinant";  
RL Arch. Virol. 144:2377-2387(1999);  
DR EMBL; AF190817; AAF13920.1; -;  
DR InterPro; IPR004005; Calici.coat.  
DR Pfam; PF00915; Calici.coat; 1.  
SQ SEQUENCE 548 AA; 53872 MW; 46DCC4042967CD79 CRC64;

Query Match 98.4%; Score 2851; DB 12; Length 548;  
Best Local Similarity 98.2%; Pred. No. 1.9e-217;  
Matches 538; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
QY 1 MKMASNDAPSNDAAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWINNFVQAP 60  
Db 1 MKMASNDATPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWINNFVQAP 60  
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHARMYNGYAGGFVQVVLGNATAGKII 120  
Db 61 GGEFTVSPRNPGEVLLNLELGPENPYLAHARMYNGYAGGFVQVVLGNATAGKII 120  
QY 121 FAAPNPFIDNLSAAQITMCPHVIVDVRLQEPVNLPMFVDRNNFFHYNQSDSRLRLIA 180  
Db 121 FAAPNPFIDNLSAAQITMCPHVIVDVRLQEPVNLPMFVDRNNFFHYNQSDSRLRLIA 180

QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPPTVESKTKPFTLPILITISEMSNS 240  
Db 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPPTVESKTKPFTLPILITISEMSNS 240  
QY 241 RFPVPISLHTSPTENIVVQCONGRVTLDELGMGTQQLPSQICAFRGVLTRSTRASDQ 300  
Db 241 RFPVPISLHTSPTENIVVQCONGRVTLDELGMGTQQLPSQICAFRGVLTRSTRASDQ 300  
QY 301 ADTATPRLFNYYWHVQDNLNGTTPYDPAEDIPGLGTPDFRGKVFQVASQORNLDSSTTRAH 360  
Db 301 ADTATPRLFNYYWHVQDNLNGTTPYDPAEDIPGLGTPDFRGKVFQVASQORNLDSSTTRAH 360  
QY 361 EAKVDTTACGFTPKLGSLEISTDSDDFQDQNTKFTTPVGIGVDNEAEFQWMLSDPYSGOF 420  
Db 361 EAKVDTTACGFTPKLGSLEISTDSDDFQDQNTKFTTPVGIGVDNEAEFQWMLSDPYSGOF 420  
QY 421 THNNMLAPAVAPNFPGEQLLFFRSQLPSSGGRSNGVLDCLVQEWVQHFYQESAPAQTOV 480  
Db 421 THNNMLAPAVAPNFPGEQLLFFRSQLPSSGGRSNGVLDCLVQEWVQHFYQESAPAQTOV 480  
QY 481 ALVRYNPDGTGKVLFEAKLHKLGFMTIANNQSDSPITVPPNGYFRFESWNPFTYTLAPMGT 540  
Db 481 ALVRYNPDGTGKVLFEAKLHKLGFMTIANNQSDSPITVPPNGYFRFESWNPFTYTLAPMGT 540  
QY 541 GNGRRRIQ 548  
Db 541 GNGRRRIQ 548

RESULT 3  
Q916E5 PRELIMINARY; PRT; 548 AA.  
ID Q916E5  
AC Q916E5;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus NLV/Oberhausen 455/01/DE.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OC NCBI\_TaxID=173921;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NLV/Oberhausen 455/01/DE;  
RA Kuenkel U., Hoehne M., Schreier E.;  
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF425768; AAL18861.1; -;  
DR InterPro; IPR004005; Calici.coat.  
DR InterPro; IPR008975; Viral\_cap.coat.  
DR Pfam; PF00915; Calici.coat; 1.  
SQ SEQUENCE 548 AA; 55958 MW; 8BA72E4DFD0330D0 CRC64;

Query Match 98.4%; Score 2850; DB 12; Length 548;  
Best Local Similarity 97.6%; Pred. No. 2.3e-217;  
Matches 535; Conservative 10; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MKMASNDAPSNDAAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWINNFVQAP 60  
Db 1 MKMASNDATPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWINNFVQAP 60  
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHARMYNGYAGGFVQVVLGNATAGKII 120  
Db 61 GGEFTVSPRNPGEVLLNLELGPENPYLAHARMYNGYAGGFVQVVLGNATAGKII 120  
QY 121 FAAPNPFIDNLSAAQITMCPHVIVDVRLQEPVNLPMFVDRNNFFHYNQSDSRLRLIA 180  
Db 121 FAAPNPFIDNLSAAQITMCPHVIVDVRLQEPVNLPMFVDRNNFFHYNQSDSRLRLIA 180  
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPPTVESKTKPFTLPILITISEMSNS 240  
Db 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPPTVESKTKPFTLPILITISEMSNS 240

```
QY 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTSTSRASDQ 300
DB 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTSTSRASDQ 300
QY 301 ADTATPRLFNYYWHVQVOLDNLNGTVPYDPAEDIPGLGTDPFRGKVFQVASQORNLDSTTRAH 360
DB 301 ADTATPRLFNYYWHVQVOLDNLNGTVPYDPAEDIPGLGTDPFRGKVFQVASQORNLDSTTRAH 360
QY 361 EAKVDTTACGRFTPKLGSLEISTSDDFDQONQTKFTVPVIGVGDNEAEFQWLSLPDYSQGF 420
DB 361 EAKVDTTACGRFTPKLGSLEISTSDDFDQONQTKFTVPVIGVGDNEAEFQWLSLPDYSQGF 420
QY 421 THNNMLAPAVANFPFGEQQLLFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTOV 480
DB 421 THNNMLAPAVANFPFGEQQLLFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTOV 480
QY 481 ALRVYVNPDTGKVLFEAKLHKGFTMTIANGDSPITVPNGYFRFESWVNPFTYLA PMGT 540
DB 481 ALRVYVNPDTGKVLFEAKLHKGFTMTIANGDSPITVPNGYFRFESWVNPFTYLA PMGT 540
QY 541 GNGRRRIQ 548
DB 541 GNGRRRIQ 548
```

## RESULT 4

```
Q8BC90 PRELIMINARY; PRT; 548 AA.
AC Q8BC90;
DT 01-MAR-2003 (TremBLrel. 23, Created)
DT 01-MAR-2003 (TremBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Herzberg 385/01/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_TaxID=207828;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Herzberg 385/01/DE;
RT "Molecular epidemiology of recombinant human calicivirus circulating in
RA Kuenkel U., Schreier E.;
RT "Characterization of recombinant human calicivirus circulating in
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF539439; AAN16370.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59954 MW; D0C02A8E89D6BBD6 CRC64;
```

Query Match 97.9%; Score 2836; DB 12; Length 548;

Best Local Similarity 97.3%; Pred. No. 3e-216; Mismatches 10; Indels 0; Gaps 0;

Matches 533; Conservative 10; Mismatches 5; Indels 0; Gaps 0;

```
QY 1 MKMASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIDPWIMNMFVQAP 60
DB 1 MKMASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIDPWIMNMFVQAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPINPYLAHARMYNGVAGFEVQVVLGNAGTAAGKII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGPINPYLAHARMYNGVAGFEVQVVLGNAGTAAGKII 120
QY 121 FFAIPNPFIDNLAAQITMCPHVIIVDVRQLEPVLNMPDVRNPNFHYNQGSDSLRLIA 180
DB 121 FFAIPNPFIDNLAAQITMCPHVIIVDVRQLEPVLNMPDVRNPNFHYNQGSDSLRLIA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTPFTLPILTISEMSNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTPFTLPILTISEMSNS 240
QY 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTSTSRASDQ 300
```

```
DB 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTSTSRASDQ 300
QY 301 ADTATPRLFNYYWHVQVOLDNLNGTVPYDPAEDIPGLGTDPFRGKVFQVASQORNLDSTTRAH 360
DB 301 ADTATPRLFNYYWHVQVOLDNLNGTVPYDPAEDIPGLGTDPFRGKVFQVASQORNLDSTTRAH 360
QY 361 EAKVDTTACGRFTPKLGSLEISTSDDFDQONQTKFTVPVIGVGDNEAEFQWLSLPDYSQGF 420
DB 361 EAKVDTTACGRFTPKLGSLEISTSDDFDQONQTKFTVPVIGVGDNEAEFQWLSLPDYSQGF 420
QY 421 THNNMLAPAVANFPFGEQQLLFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTOV 480
DB 421 THNNMLAPAVANFPFGEQQLLFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTOV 480
QY 481 ALRVYVNPDTGKVLFEAKLHKGFTMTIANGDSPITVPNGYFRFESWVNPFTYLA PMGT 540
DB 481 ALRVYVNPDTGKVLFEAKLHKGFTMTIANGDSPITVPNGYFRFESWVNPFTYLA PMGT 540
QY 541 GNGRRRIQ 548
DB 541 GNGRRRIQ 548
```

## RESULT 5

```
Q915D2 PRELIMINARY; PRT; 548 AA.
AC Q915D2;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Bitburg/289/01/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_TaxID=173924;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Bitburg/289/01/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RT with Norwalk-like viruses in Germany.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF427112; AAL18864.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59823 MW; A83BC58F87F93088 CRC64;
```

Query Match 97.8%; Score 2831; DB 12; Length 548;

Best Local Similarity 97.4%; Pred. No. 7.5e-216; Mismatches 7; Indels 0; Gaps 0;

Matches 534; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

```
QY 1 MKMASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIDPWIMNMFVQAP 60
DB 1 MKMASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIDPWIMNMFVQAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPINPYLAHARMYNGVAGFEVQVVLGNAGTAAGKII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGPINPYLAHARMYNGVAGFEVQVVLGNAGTAAGKII 120
QY 121 FFAIPNPFIDNLAAQITMCPHVIIVDVRQLEPVLNMPDVRNPNFHYNQGSDSLRLIA 180
DB 121 FFAIPNPFIDNLAAQITMCPHVIIVDVRQLEPVLNMPDVRNPNFHYNQGSDSLRLIA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTPFTLPILTISEMSNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTPFTLPILTISEMSNS 240
QY 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTSTSRASDQ 300
DB 241 RPPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTSTSRASDQ 300
QY 301 ADTATPRLFNYYWHVQVOLDNLNGTVPYDPAEDIPGLGTDPFRGKVFQVASQORNLDSTTRAH 360
```

```
Db 301 ADTATPRLFNYYWHIQLONLNGTDCDPAEDIPGLTDFGRKVGFGVASQRNPSTTGAH 360
QY 361 EAKVDTTAGRTPKLGSLEISTSDDFDQONQTKTTPVGIQVDNEAEFQOWSLPDYSQGF 420
Db 361 EAKVDTTAGRTPKLGSLEISTSDDFDQONQTKTTPVGIQVDNEAEFQOWSLPDYSQGF 420
QY 421 THNMNLAPAVAPNPFQEQQLLFFRSQLPSSGGRSNGVLDCLVPOEWVQHFYQESAPAOQV 480
Db 421 THNMNLAPAVAPNPFQEQQLLFFRSQLPSSGGRSNGVLDCLVPOEWVQHFYQESAPAOQV 480
QY 481 ALVRYNPDGTGKVLFEAKLHKLGFMTIANNQDSPTVPPNGYFRFESWNPFTYLLAPMGT 540
Db 481 ALVRYNPDGTGKVLFEAKLHKLGFMTIANNQDSPTVPPNGYFRFESWNPFTYLLAPMGT 540
QY 541 GNGRRRIQ 548
Db 541 GNGRRRIQ 548

RESULT 6
Q66296 PRELIMINARY; PRT; 548 AA.
ID Q66296
AC Q66296;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLIV/Berlin/226/01/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae.
OC* Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae.
OX NCBI_TaxID=11975;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=TV24;
RX MEDLINE=94202329; PubMed=8151799;
RA Lew J.F., Petric M., Kapikian A.Z., Jiang X., Estes M.K., Green K.Y.;
RT "Identification of Minireovirus as a Norwalk-like virus in pediatric patients with gastroenteritis."
RL J. Virol. 68:3391-3396(1994).
DR EMBL; U02030; AAA18930.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 60034 MW; CCEE4A180E85008B CRC64;

Query Match 97.4%; Score 2822; DB 12; Length 548;
Best Local Similarity 96.7%; Pred. No. 3.9e-215;
Matches 530; Conservative 11; Mismatches 7; Indels 0; Gaps 0;

QY 1 MKMASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNMFVQAP 60
Db 1 MKMASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNMFVQAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPINPYLAHLARMYNGYAGGFEVQVVLGNAGTAKGII 120
Db 61 GGEFTVSPRNSPGEVLLNLELGPINPYLAHLARMYNGYAGGFEVQVVLGNAGTAKGII 120
QY 121 FAALPPNPIDNLSAAQITMCPHVIVDVRQLEPNLPMQVDRNFFHYNQSGSRLRLIA 180
Db 121 FAALPPNPIDNLSAAQITMCPHVIVDVRQLEPNLPMQVDRNFFHYNQSGSRLRLIA 180
QY 181 MLYTPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPVTSKTKPFTLPILITISEMSNS 240
Db 181 MLYTPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPVTSKTKPFTLPILITISEMSNS 240
QY 241 RFPVPIESLHSTPTENIVVQCONGRVTLDELMGTQTLPSQICAFRGVLRSTRSRASDQ 300
Db 241 RFPVPIESLHSTPTENIVVQCONGRVTLDELMGTQTLPSQICAFRGVLRSTRSRASDQ 300
QY 301 ADTATPRLFNYYWHIQLONLNGTDCDPAEDIPGLTDFGRKVGFGVASQRNPSTTGAH 360
Db 301 ADTATPRLFNYYWHIQLONLNGTDCDPAEDIPGLTDFGRKVGFGVASQRNPSTTGAH 360
```

```
QY 361 EAKVDTTAGRTPKLGSLEISTSDDFDQONQTKTTPVGIQVDNEAEFQOWSLPDYSQGF 420
Db 361 EAKVDTTAGRTPKLGSLEISTSDDFDQONQTKTTPVGIQVDNEAEFQOWSLPDYSQGF 420
QY 421 THNMNLAPAVAPNPFQEQQLLFFRSQLPSSGGRSNGVLDCLVPOEWVQHFYQESAPAOQV 480
Db 421 THNMNLAPAVAPNPFQEQQLLFFRSQLPSSGGRSNGVLDCLVPOEWVQHFYQESAPAOQV 480
QY 481 ALVRYNPDGTGKVLFEAKLHKLGFMTIANNQDSPTVPPNGYFRFESWNPFTYLLAPMGT 540
Db 481 ALVRYNPDGTGKVLFEAKLHKLGFMTIANNQDSPTVPPNGYFRFESWNPFTYLLAPMGT 540
QY 541 GNGRRRIQ 548
Db 541 GNGRRRIQ 548

RESULT 7
Q915D3 PRELIMINARY; PRT; 548 AA.
ID Q915D3
AC Q915D3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLIV/Berlin/226/01/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus
OX NCBI_TaxID=173923;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLIV/Berlin/226/01/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany."
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF427111; AALI8863.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59844 MW; 391B68558ABDFE85 CRC64;

Query Match 97.3%; Score 2819; DB 12; Length 548;
Best Local Similarity 96.9%; Pred. No. 6.7e-215;
Matches 531; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKMASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNMFVQAP 60
Db 1 MKMASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNMFVQAP 60
QY 61 GGEFTVSPRNSPGEVLLNLELGPINPYLAHLARMYNGYAGGFEVQVVLGNAGTAKGII 120
Db 61 GGEFTVSPRNSPGEVLLNLELGPINPYLAHLARMYNGYAGGFEVQVVLGNAGTAKGII 120
QY 121 FAALPPNPIDNLSAAQITMCPHVIVDVRQLEPNLPMQVDRNFFHYNQSGSRLRLIA 180
Db 121 FAALPPNPIDNLSAAQITMCPHVIVDVRQLEPNLPMQVDRNFFHYNQSGSRLRLIA 180
QY 181 MLYTPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPVTSKTKPFTLPILITISEMSNS 240
Db 181 MLYTPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPVTSKTKPFTLPILITISEMSNS 240
QY 241 RFPVPIESLHSTPTENIVVQCONGRVTLDELMGTQTLPSQICAFRGVLRSTRSRASDQ 300
Db 241 RFPVPIESLHSTPTENIVVQCONGRVTLDELMGTQTLPSQICAFRGVLRSTRSRASDQ 300
QY 301 ADTATPRLFNYYWHIQLONLNGTDCDPAEDIPGLTDFGRKVGFGVASQRNPSTTGAH 360
Db 301 ADTATPRLFNYYWHIQLONLNGTDCDPAEDIPGLTDFGRKVGFGVASQRNPSTTGAH 360
QY 361 EAKVDTTAGRTPKLGSLEISTSDDFDQONQTKTTPVGIQVDNEAEFQOWSLPDYSQGF 420
Db 361 EAKVDTTAGRTPKLGSLEISTSDDFDQONQTKTTPVGIQVDNEAEFQOWSLPDYSQGF 420
```

```
QY 421 THNNMLAPAVANFPGEQLLFRSOLPSSGGRSNGVLDCLVPOEWWOHPYQESAPAQTOV 480
Db 421 THNNMLAPAVANFPGEQLLFRSOLPSSGGRSNGVLDCLVPOEWWOHPYQESAPAQTOV 480
QY 481 ALVRVYNPDGTGVLFEAKLHKGFTMTIANNGDSPITVPNGYFRFESWNPVYTLAPMGT 540
Db 481 ALVRVYNPDGTGVLFEAKLHKGFTMTIANNGDSPITVPNGYFRFESWNPVYTLAPMGT 540
QY 541 GNGRRRIQ 548
Db 541 GNGRRRIQ 548

RESULT 8
QVOP5
ID Q8VOP5 PRELIMINARY; PRT; 548 AA.
AC Q8VOP5;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus Hu/NLV/GII/MD101-2/1987/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=159311;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/GII/MD101-2/1987/US;
RX MEDLINE=21666333; PubMed=11807686;
RA Green K.Y., Belliot G., Taylor J.L., Valdesuso J., Lew J.F.,
RA Kapikian A.Z., Lin F.Y.C.;
RT "A Predominant Role for Norwalk-like Viruses as Agents of Epidemic
RT Gastroenteritis in Maryland Nursing Homes for the Elderly.";
RL J. Infect. Dis. 185:133-146(2002).
DR EMBL: AY030312; RAK54360.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59898 MW; 776F74B4E58B6CC7 CRC64;
```

```
Query Match 97.2%; Score 2815; DB 12; Length 548;
Best Local Similarity 96.2%; Pred. No. 1.4e-214;
Matches 527; Conservative 15; Mismatches 6; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQOIIIDPWIMNPFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQOIIIDPWIMNPFVQAP 60
QY 61 GGEFTVSPNSPGEVLLNLLELGEINPNYLAHLARMYNGYAGGFVQVVLGNAFTAGKII 120
Db 61 GGEFTVSPNSPGEVLLNLLELGEINPNYLAHLARMYNGYAGGFVQVVLGNAFTAGKVI 120
QY 121 FFAAIPNPFIDNLSAAQITMCCHVIVDVRQLEPVLNPDVNRNFFHYNQGSDSLRLIA 180
Db 121 FFAAIPNPFIDNLSAAQITMCCHVIVDVRQLEPVLNPDVNRNFFHYNQGSDSLRLIA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTPSPDFSNFLVPPVTESKTPFTLPIITISEMNS 240
Db 181 MLYTPLRANSGDDVFTVSCRVLTPSPDFSNFLVPPVTESKTPFTLPIITISEMNS 240
QY 241 RFPVPVIESLHTSPENIVVQCONGRVTLTGELMGTTQLLPISOICAFRGVLTSTGRASDQ 300
Db 241 RFPVPVIESLHTSPENIVVQCONGRVTLTGELMGTTQLLPISOICAFRGVLTSTGRASDQ 300
QY 301 ADTATPRLENYVHVLQDNLNGTTPDPAEDIPGLGTPDFRGKVFVGASQRNLDTSTTRAH 360
Db 301 ADTATPRLENYVHVLQDNLNGTTPDPAEDIPGLGTPDFRGKVFVGASQRNLDTSTTRAH 360
QY 361 EAKVDTTAGRTFKLGSLEISDSDFDQONQPTKFTPVGIGVDNEAEFOQWSLPDYSQGF 420
Db 361 EAKVDTTAGRTFKLGSLEISDSDFDQONQPTKFTPVGIGVDNEAEFOQWSLPDYSQGF 420
```

```
QY 421 THNNMLAPAVANFPGEQLLFRSOLPSSGGRSNGVLDCLVPOEWWOHPYQESAPAQTOV 480
Db 421 THNNMLAPAVANFPGEQLLFRSOLPSSGGRSNGVLDCLVPOEWWOHPYQESAPAQTOV 480
QY 481 ALVRVYNPDGTGVLFEAKLHKGFTMTIANNGDSPITVPNGYFRFESWNPVYTLAPMGT 540
Db 481 ALVRVYNPDGTGVLFEAKLHKGFTMTIANNGDSPITVPNGYFRFESWNPVYTLAPMGT 540
QY 541 GNGRRRIQ 548
Db 541 GNGRRRIQ 548

RESULT 9
QVOP5
ID Q918A1 PRELIMINARY; PRT; 548 AA.
AC Q918A1;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein.
OS Norwalk-like virus NLV/New Orleans/279/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171844;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RT Submitted (AUG-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL: AF414412; AAL2989.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59990 MW; F998C603777P8595 CRC64;
```

```
Query Match 97.1%; Score 2812; DB 12; Length 548;
Best Local Similarity 96.4%; Pred. No. 2.4e-214;
Matches 528; Conservative 12; Mismatches 8; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQOIIIDPWIMNPFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQOIIIDPWIMNPFVQAP 60
QY 61 GGEFTVSPNSPGEVLLNLLELGEINPNYLAHLARMYNGYAGGFVQVVLGNAFTAGKII 120
```

Db	61	:     :
----	----	---

Db	61	CGEFTVSPRNSPCEVLLNLELGPENIPYLAHLARMYNGYAGGFVEQVVLGNAGNAFTAGKVI	120
Qy	121	FAAIPNPFIDNLSAAQITMCCHVIVDVROLEPVLNMPDVRNPFHYNQSDSRRLRIA	180
Db	121	FAAIPNPFIDNLSAAQITMCCHVIVDVROLEPVLNMPDVRNPFHYNQSDSRRLRIA	180
Qy	181	MLYTPPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTKPFTLPILTISEMSNS	240
Db	181	MLYTPPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTKPFTLPILTISEMSNS	240
Qy	241	RPVPVIESLHTSPENIVVOCONGRVTLTGELMGTTQLLPSCICAFRGVLTSTSRASDQ	300
Db	241	RPVPVIESLHTSPENIVVOCONGRVTLTGELMGTTQLLPSCICAFRGVLTSTSRASDQ	300
Qy	301	ADTATPRLFNYYVHVDNLNGTYPDAEDIIPGLTDPDFRGKVGVSQRNLDSTTRAH	360
Db	301	ADTATPRLFNYYVHVDNLNGTYPDAEDIIPGLTDPDFRGKVGVSQRNLDSTTRAH	360
Qy	361	EAKVDTTAGRTFKLGSLEISTSDDDFQNPQTKFTFPGVIGVDNEAFQOWSLPDYSGOF	420
Db	361	EAKVDTTAGRTFKLGSLEISTSDDDFQNPQTKFTFPGVIGVDNEAFQOWSLPDYSGOF	420
Qy	421	THNNLAPAVAPNFPGEQLLFFRSQLPSSGSRNGVLDCLVPOEWQHFYQESAPACTOV	480
Db	421	THNNLAPAVAPNFPGEQLLFFRSQLPSSGSRNGVLDCLVPOEWQHFYQESAPACTOV	480
Qy	481	ALVRYVNPDTGKVLFEAKLHKLGFMTIANGSDSPITVPPNGYFRFESWVNPFTYLPMTG	540
Db	481	ALVRYVNPDTGKVLFEAKLHKLGFMTIANGSDSPITVPPNGYFRFESWVNPFTYLPMTG	540
Qy	541	NGRRRIQ 548	
Db	541	NGRRRIQ 548	
RESULT 11			
Q8V768	PRELIMINARY; PRT; 548 AA.		
AC	Q8V768		
DT	01-MAR-2002 (Tremblrel. 20, Created)		
DT	01-MAR-2002 (Tremblrel. 20, Last sequence update)		
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)		
DE	ORF2 protein (Capsid).		
OS	Norwalk virus, and		
OS	Norwalk-like virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_TaxID=11983, 95340;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	SPECIES=Norwalk virus; STRAIN=U201GII;		
RA	Kojima S., Kageyama T., Fukushi S., Hoshino F.B., Shinohara M.,		
RA	Uchida K., Natori K., Takeda N., Katayama K.;		
RT	"Genogroup-specific primers for detect Norwalk like virus.;"		
RL	Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	SPECIES=Norwalk-like virus; STRAIN=Saitama U201;		
RA	Katayama K.;		
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	SPECIES=Norwalk-like virus; STRAIN=Saitama U201;		
RX	MEDLINE=22192455; PubMed=12202225;		
RA	Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,		
RA	Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,		
RA	Gojobori T., Takeda N.;		
RT	"Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like		
RT	Viruses.;"		
RL	Virology 299:225-239 (2002).		
DR	EMBL; AB067542; BAB84155.1; -.		
DR	EMBL; AB039782; BAC11837.1; -.		
DR	InterPro; IPR004005; Calici_coat.		



DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 548 AA; 59949 MW; 589FC65F14B80F64 CRC64;

Query Match 96.8%; Score 2802; DB 12; Length 548;  
Best Local Similarity 96.4%; Pred. No. 1.5e-213;  
Matches 528; Conservative 11; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIDPWIMNNFVQAP 60  
DB 1 MKMASNDAAPSNDGAAGLVPEINNEAMALEPVAGAAIAAPLTGQQNIIDPWIMNNFVQAP 60

QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGFEVQVVLGNAGNFTAGKII 120  
DB 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGFEVQVVLGNAGNFTAGKII 120

QY 121 FFAIIPNPFIIDNLSAAQITMCPHVIVDVRLPEVNLPMPPDVNRNFFHYNQGSDSLRLIA 180  
DB 121 FFAIIPNPFIIDNLSAAQITMCPHVIVDVRLPEVNLPMPPDVNRNFFHYNQGSDSLRLIA 180

QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSFNFLVPPPTVESKTKPFTLPILTISEMSNS 240  
DB 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSFNFLVPPPTVESKTKPFTLPILTISEMSNS 240

QY 241 RFPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTTRSTGRASDQ 300  
DB 241 RFPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTTRSTGRASDQ 300

QY 241 RFPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTTRSTGRASDQ 300  
DB 241 RFPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTTRSTGRASDQ 300

QY 301 ADTATPRLFNYYWHVQLDNLNGTVPYPAEDIPGLGTPDFRGKVGVSQORNLDSSTTRAH 360  
DB 301 ADTATPRLFNYYWHVQLDNLNGTVPYPAEDIPGLGTPDFRGKVGVSQORNLDSSTTRAH 360

QY 361 EAKVDTTACGRTPKLGSLSEISTDSDDFQONQPTKFTPVGIGVDNEAEFQOWSLPDYSGOF 420  
DB 361 EAKVDTTACGRTPKLGSLSEISTDSDDFQONQPTKFTPVGIGVDNEAEFQOWSLPDYSGOF 420

QY 421 THNMNLAPAVAPNPFGEQLLFFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQOV 480  
DB 421 THNMNLAPAVAPNPFGEQLLFFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQOV 480

QY 481 ALVRYVNPDTGKVLPEAKLHKLGMFTIANNGDSPITVPNGYFRFESVWNPFTYLPAMGT 540  
DB 481 ALVRYVNPDTGKVLPEAKLHKLGMFTIANNGDSPITVPNGYFRFESVWNPFTYLPAMGT 540

QY 541 GNGRRRIQ 548  
DB 541 GNGRRRIQ 548

RESULT 12  
Q88291 ID Q88291 PRELIMINARY; PRT; 548 AA.  
AC Q88291  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Small round structured virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=37141;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Auckland;  
RA Regi W.J., Green D.H., Lewis G.D.;  
RT "Outbreaks of oyster associated gastroenteritis: Virological  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U46039; AAB00437.1; -  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 548 AA; 59907 MW; 364BFA6EC263B6B2 CRC64;

Query Match 96.7%; Score 2799; DB 12; Length 548;  
Best Local Similarity 95.6%; Pred. No. 2.6e-213;  
Matches 524; Conservative 15; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQQNIIDPWIMNNFVQAP 60  
DB 1 MKMASNDAAPSNDGAAGLVPEINNEAMALEPVAGAAIAAPLTGQQNIIDPWIMNNFVQAP 60

QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGFEVQVVLGNAGNFTAGKII 120  
DB 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGFEVQVVLGNAGNFTAGKII 120

QY 121 FFAIIPNPFIIDNLSAAQITMCPHVIVDVRLPEVNLPMPPDVNRNFFHYNQGSDSLRLIA 180  
DB 121 FFAIIPNPFIIDNLSAAQITMCPHVIVDVRLPEVNLPMPPDVNRNFFHYNQGSDSLRLIA 180

QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSFNFLVPPPTVESKTKPFTLPILTISEMSNS 240  
DB 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSFNFLVPPPTVESKTKPFTLPILTISEMSNS 240

QY 241 RFPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTTRSTGRASDQ 300  
DB 241 RFPVPIESLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTTRSTGRASDQ 300

QY 301 ADTATPRLFNYYWHVQLDNLNGTVPYPAEDIPGLGTPDFRGKVGVSQORNLDSSTTRAH 360  
DB 301 ADTATPRLFNYYWHVQLDNLNGTVPYPAEDIPGLGTPDFRGKVGVSQORNLDSSTTRAH 360

QY 361 EAKVDTTACGRTPKLGSLSEISTDSDDFQONQPTKFTPVGIGVDNEAEFQOWSLPDYSGOF 420  
DB 361 EAKVDTTACGRTPKLGSLSEISTDSDDFQONQPTKFTPVGIGVDNEAEFQOWSLPDYSGOF 420

QY 421 THNMNLAPAVAPNPFGEQLLFFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQOV 480  
DB 421 THNMNLAPAVAPNPFGEQLLFFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQOV 480

QY 481 ALVRYVNPDTGKVLPEAKLHKLGMFTIANNGDSPITVPNGYFRFESVWNPFTYLPAMGT 540  
DB 481 ALVRYVNPDTGKVLPEAKLHKLGMFTIANNGDSPITVPNGYFRFESVWNPFTYLPAMGT 540

QY 541 GNGRRRIQ 548  
DB 541 GNGRRRIQ 548

RESULT 13  
Q83880 ID Q83880 PRELIMINARY; PRT; 548 AA.  
AC Q83880;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
DE Capsid protein (Fragment).  
OS Norwalk virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=11983;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=SRV-OTH-25/89/J;  
RX MEDLINE=94335115; PubMed=8057474;  
RA Wang J., Jiang X., Madore H.P., Gray J., Desselberger U., Ando T.,  
RA Seto Y., Oishi I., Lew J.F., Green K.Y., et al.;  
RT "Sequence diversity of small, round-structured viruses in the Norwalk  
RL virus group."  
RL J. Virol. 68:5982-5990 (1994).  
DR EMBL: L23830; AAA59233.1; -  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
FT NON\_TER 548  
SQ SEQUENCE 548 AA; 59904 MW; 054379C8AA7541D9 CRC64;

RA Seto Y., Monroe S.S., Glass R.I.;  
RT "Correlation of patient immune responses with genetically  
RT characterized small round-structured viruses involved in outbreaks of  
RT nonbacterial acute gastroenteritis in the United States, 1990 to  
RT 1995." J. Med. Virol. 53:372-383 (1997).  
RL [3]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Hu/NLV/Towson/313/1994/US;  
RC MEDLINE=20266011; PubMed=10804147;  
RX Ando T., Noel J.S., Fankhauser R.L.;  
RA "Genetic classification of 'Norwalk-like viruses.'";  
RT J. Infect. Dis. 181:S336-S348 (2000).  
RL [4]  
RN SEQUENCE FROM N.A.  
RP STRAIN=Hu/NLV/Towson/313/1994/US;  
RC Ando T., Seto Y., Noel J.S., Monroe S.S., Glaes R.I., Fankhauser R.L.;  
RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF414414; AAL12995.1; -;  
DR InterPro; IPR004005; Calici.coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calici\_coat.1.  
SQ SEQUENCE 548 AA; 5993 MW; 114F3907B3A26D89 CRC64;

Query Match 96.2%; Score 2786; DB 12; Length 548;  
Best Local Similarity 96.0%; Pred. No. 2.8e-212;  
Matches 526; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWIMNFFVQAP 60  
Db 1 MKMASNDAAPSNDGAAGLVPEINNEAMALEPVGAAIAAPLTGQNIIDPWIMNFFVQAP 60  
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFVQVVLVAGNATAGKII 120  
Db 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFVQVVLVAGNATAGKVI 120  
QY 121 FAAPPNPPIIDLSAAQITMCPHVIVDVRLQEPINLPMPPDVVRNPFHYNQSDSLRLIA 180  
Db 121 FAAPPNPPIIDLSAAQITMCPHVIVDVRLQEPINLPMPPDVVRNPFHYNQSDSLRLIA 180  
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSPDFSNFLVPSTMESKTPFTLPILTISEMNS 240  
Db 181 MLYTPLRANSGDDVFTVSCRVLTRPSPDFSNFLVPSTMESKTPFTLPILTISEMNS 240  
QY 241 RPPVPISLHTSPTENIVVQCGRVTLDGELMGTTQLLPNQICAFRGTLTRSTNRASDQ 300  
Db 241 RPPVPIDSLHTSPTENIVVQCGRVTLDGELMGTTQLLPNQICAFRGTLTRSTNRASDQ 300  
QY 301 ADTATPRLFNYYHVLQDNLNGTTPYDPAEDIPGPIGTDPFRGKVGVSQRNLDSTTRAH 360  
Db 301 ADTATPRLFNYYHVLQDNLNGTTPYDPAEDIPGPIGTDPFRGKVGVSQRNLDSTTRAH 360  
QY 361 EAKVDTTAGRFTPKLGSLEISTDSDDFDQNPQKFTFVIGIVDNEAEFQOWSLPDYSGOF 420  
Db 361 EAKVDTTSGRFTPKLGSLEIITSDSDFDTNQSTKFTFVIGIVDNEAEFQOWSLPDYSGOF 420  
QY 421 THNMNLAPAVANFPFGEQLLFFRSQIPSSGGRSNGVLDCLVPOEWVQHFYQSSAPAQTOV 480  
Db 421 THNMNLAPAVANFPFGEQLLFFRSQIPSSGGRSNGVLDCLVPOEWVQHFYQSSAPAQTOV 480  
QY 481 ALVRYVNPDTGRVLFEAKLHKLGFMTIANGDSPTIVPENGYFRFESWNPFTYLPAMGT 540  
Db 481 ALVRYVNPDTGRVLFEAKLHKLGFMTIANGDSPTIVPENGYFRFESWNPFTYLPAMGT 540  
QY 541 GNGRRRIQ 548  
Db 541 GNGRRRIQ 548

RESULT 15  
Q91722 PRELIMINARY; PRT; 548 AA.  
ID Q91722  
AC Q91722;

Query Match 96.5%; Score 2794; DB 12; Length 548;  
Best Local Similarity 95.4%; Pred. No. 6.4e-213;  
Matches 523; Conservative 16; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWIMNFFVQAP 60  
Db 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPWIMNFFVQAP 60  
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFVQVVLVAGNATAGKII 120  
Db 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFVQVVLVAGNATAGKVI 120  
QY 121 FAAPPNPPIIDLSAAQITMCPHVIVDVRLQEPINLPMPPDVVRNPFHYNQSDSLRLIA 180  
Db 121 FAAPPNPPIIDLSAAQITMCPHVIVDVRLQEPINLPMPPDVVRNPFHYNQSDSLRLIA 180  
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSPDFSNFLVPSTMESKTPFTLPILTISEMNS 240  
Db 181 MLYTPLRANSGDDVFTVSCRVLTRPSPDFSNFLVPSTMESKTPFTLPILTISEMNS 240  
QY 241 RPPVPISLHTSPTENIVVQCGRVTLDGELMGTTQLLPNQICAFRGTLTRSTNRASDQ 300  
Db 241 RPPVPIDSLHTSPTENIVVQCGRVTLDGELMGTTQLLPNQICAFRGTLTRSTNRASDQ 300  
QY 301 ADTATPRLFNYYHVLQDNLNGTTPYDPAEDIPGPIGTDPFRGKVGVSQRNLDSTTRAH 360  
Db 301 ADTATPRLFNYYHVLQDNLNGTTPYDPAEDIPGPIGTDPFRGKVGVSQRNLDSTTRAH 360  
QY 361 EAKVDTTAGRFTPKLGSLEISTDSDDFDQNPQKFTFVIGIVDNEAEFQOWSLPDYSGOF 420  
Db 361 EAKVDTTSGRFTPKLGSLEIITSDSDFDTNQSTKFTFVIGIVDNEAEFQOWSLPDYSGOF 420  
QY 421 THNMNLAPAVANFPFGEQLLFFRSQIPSSGGRSNGVLDCLVPOEWVQHFYQSSAPAQTOV 480  
Db 421 THNMNLAPAVANFPFGEQLLFFRSQIPSSGGRSNGVLDCLVPOEWVQHFYQSSAPAQTOV 480  
QY 481 ALVRYVNPDTGRVLFEAKLHKLGFMTIANGDSPTIVPENGYFRFESWNPFTYLPAMGT 540  
Db 481 ALVRYVNPDTGRVLFEAKLHKLGFMTIANGDSPTIVPENGYFRFESWNPFTYLPAMGT 540  
QY 541 GNGRRRIQ 548  
Db 541 GNGRRRIQ 548

RESULT 14  
Q91725 PRELIMINARY; PRT; 548 AA.  
ID Q91725  
AC Q91725  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
GN ORF2.  
OS Norwalk-like virus NLV/Towson/313/1994/US.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=171846;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Towson/313/1994/US;  
RX MEDLINE=97193806; PubMed=9041391;  
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
RT of small round-structured viruses (Norwalk-like viruses).";  
RL J. Clin. Microbiol. 35:570-577 (1997).  
RN [2]  
RP SEQUENCE OF 100-192 FROM N.A.  
RC STRAIN=Hu/NLV/Towson/313/1994/US;  
RX MEDLINE=98071277; PubMed=9407386;  
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,

DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Capsid protein.  
GN ORF2.  
OS Norwalk-like virus NLV/Brattleboro/321/1995/US.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=171847;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;  
RX MEDLINE=97193306; PubMed=9041391;  
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
RT of small round-structured viruses (Norwalk-like viruses).";  
RL J. Clin. Microbiol. 35:570-577(1997).  
RN [2]  
RP SEQUENCE OF 100-192 FROM N.A.  
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;  
RX MEDLINE=98071277; PubMed=9407386;  
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,  
RA Seto Y., Monroe S.S., Glass R.I.;  
RT "Correlation of patient immune responses with genetically  
RT characterized small round-structured viruses involved in outbreaks of  
RT nonbacterial acute gastroenteritis in the United States, 1990 to  
RT 1995.";  
RL J. Med. Virol. 53:372-383(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;  
RX MEDLINE=20266071; PubMed=10804147;  
RA Ando T., Noel J.S., Fankhauser R.L.;  
RT "Genetic classification of 'Norwalk'-like viruses.";  
RL J. Infect. Dis. 181:S336-S348(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;  
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF414415; AAL1298.1; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 548 AA; 55933 MW; 7B8AFF9AF1469158 CRC64;

Query Match 96.0%; Score 2781; DB 12; Length 548;  
Best Local Similarity 95.8%; Pred. No. 6.9e-212;  
Matches 525; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAPLTGQONIIDPWIMNNFVOAP 60  
DB 1 MKMASNDAAPSNDGAAGLVPEINNEAMALEPVAGAAIAPLTGQONIIDPWIMNNFVOAP 60  
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAFTAGKII 120  
DB 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAFTAGKVI 120  
QY 121 FAAPPPNPIDNLSAAQITMCPHVIVDVQRLEPVNLPMDPVRNPNFFHYNQGSDSLRLIA 180  
DB 121 FAAPPPNPIDNLSAAQITMCPHVIVDVQRLEPVNLPMDPVRNPNFFHYNQGSDSLRLIA 180  
QY 181 MLYTFLRANSGDDVFTVSCRVLTPSPDFSENFVLPPTVESKTKPFTLPILTISEMNS 240  
DB 181 MLYTFLRANSGDDVFTVSCRVLTPSPDFSENFVLPPTVESKTKPFTLPILTISEMNS 240  
QY 241 RPPVPIESLHSTPTENIVVQCONGRVTLGDMGTTLQLLPQICAFRGVLTSTSRASDQ 300  
DB 241 RPPVPIESLHSTPTENIVVQCONGRVTLGDMGTTLQLLPQICAFRGVLTSTSRASDQ 300  
QY 301 ADTATPRLFNYYVHVLQNLNGTTPYDPAEDIPGLGTPDFRGKVFQVASQRNLDSTTRAH 360  
DB 301 ADTATPRLFNYYVHVLQNLNGTTPYDPAEDIPGLGTPDFRGKVFQVASQRNLDSTTRAH 360

DB 301 ADTPTPRLFNHRWHIQLDNLNGTTPYDPAEDIPAPLGTPDFRGKVFQVASQRNPDSTTRAH 360  
QY 361 EAKVDTTATAGRTPKLGSLEISTDSDDFQONQPTKFTPVGIGVDNEAEFQWLSLPDYSGQF 420  
DB 361 EAKVDTTSDRPTPKLGSLEIITESGDFDNTQSTKFTPVGIGVDNEAEFQWLSLPDYSGQF 420  
QY 421 THNMNLAPAVAPNFFPGEQLLFFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTOV 480  
DB 421 THNMNLAPAVAPNFFPGEQLLFFRSQLPSSGGRSNGVLDCLVPQEWVQHFYQESAPAQTOV 480  
QY 481 ALVRYVNPDTGKVLPEAKLHKLGMWTIANNGDSPITVPNGYFRFESWVNPFTLAPMGT 540  
DB 481 ALVRYVNPDTGKVLPEAKLHKLGMWTIANNGDSPITVPNGYFRFESWVNPFTLAPMGT 540  
QY 541 GNGRRRIQ 548  
DB 541 GNGRRRIQ 548

Search completed: June 1, 2004, 13:53:17  
Job time : 32.9054 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.5738 Seconds  
(without alignments)  
4985.230 Million cell updates/sec

Title: US-09-926-799-6  
Perfect score: 2896  
Sequence: 1 MKMASNDAPSDNGAAGLVP.....VNPFYTLAPMGNGRRRIQ 548  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034	70.2	542	2 S60616	capsid protein - h
2	1947.5	67.2	539	2 S40111	capsid protein - h
3	1211	41.8	530	2 B37471	capsid protein - N
4	1123.5	38.8	546	2 B37491	major capsid prote
5	302.5	10.4	576	2 A53982	capsid protein - E
6	284.5	9.8	2344	1 RRWRH	genome polyprotein
7	281.5	9.7	702	1 A48562	coat protein - San
8	277.5	9.6	2344	2 S53199	genome polyprotein
9	276	9.5	2344	2 S64740	genome polyprotein
10	275.5	9.5	671	1 VCMWF9	coat protein - fel
11	269	9.3	703	1 C48562	coat protein - San
12	262	9.0	668	1 VCMWFF	coat protein - fel
13	261.5	9.0	668	1 VCMWFC	coat protein - fel
14	257	8.9	668	2 JQ2354	capsid protein - f
15	255	8.8	668	2 JQ2356	capsid protein - f
16	138.5	4.8	4936	2 AH2515	hypothetical prote
17	137.5	4.7	2194	1 GNNY17	genome polyprotein
18	133.5	4.6	3085	2 T00327	polyprotein - infe
19	128.5	4.4	2205	1 GNNY2W	genome polyprotein
20	126.5	4.4	2207	1 GNNY5P	genome polyprotein
21	126	4.4	1965	2 S75200	fat protein - Syne
22	125	4.3	2175	1 GNNYBE	genome polyprotein
23	124.5	4.3	940	2 DB9723	protein f39D8.1b [
24	124.5	4.3	945	2 T21998	hypothetical prote
25	121.5	4.2	929	2 A44048	genome polyprotein
26	121	4.2	2233	2 B95075	beta-galactosidase
27	120.5	4.2	943	2 JC4081	sucrase/fructanase
28	119	4.1	2332	1 GNNYF	genome polyprotein
29	118.5	4.1	1023	2 T48997	epsin-like protein

30	118.5	4.1	2183	1 GNNYB4	genome polyprotein
31	116.5	4.0	833	2 S45041	genome polyprotein
32	116.5	4.0	900	1 GNNYMW	genome polyprotein
33	116.5	4.0	2333	1 GNNY2F	genome polyprotein
34	116	4.0	1197	2 T30581	neural cell adhesi
35	116	4.0	2206	2 S03822	genome polyprotein
36	115.5	4.0	733	2 JQ1891	capsid protein - f
37	115	4.0	1474	2 F69009	probable membrane
38	114.5	4.0	818	2 T08823	structural polypro
39	114	3.9	733	2 JQ1892	capsid protein - f
40	114	3.9	2206	1 GNNY4P	genome polyprotein
41	113.5	3.9	429	2 S27924	gene Lf2 protein -
42	113	3.9	884	2 E75489	conserved hypotet
43	113	3.9	2207	2 S09553	genome polyprotein
44	111.5	3.9	2271	2 F90073	hypothetical prote
45	111	3.8	5175	2 T20992	hypothetical prote

ALIGNMENTS

RESULT 1

S60616  
capsid protein - human calicivirus (strain Melksham)

C:Species: human calicivirus

A:Variety: strain Melksham

C:Date: 23-May-1997 #sequence\_revision 23-May-1997 #text\_change 28-Jul-2000

C:Accession: S60616

GR:Green, S.M.; Lambden, P.R.; Caul, E.M.; Ahlstedt, G.R.; Gilevsky, A.J.

Variant Residues: 307, 277, 285, 4395

A:Title: Capsid diversity in small round-structured viruses: molecular characterization

A:Reference number: S60615; MUID:96136658; PMID:8533462

A:Accession: S60616

A:Molecule type: Genomic RNA

A:Residues: 1-542 <GRE>

A:Cross-references: EMBL:X81879; NID:G976077; PIDN:CAA57462.1; PID:G976079

A:Experimental source: strain Melksham

A:Note: it is uncertain whether Met-1 or Met-3 is the Initiator

C:Superfamily: human calicivirus capsid protein

C:Keywords: capsid protein; coat protein

Query Match 70.2%; Score 2034; DB 2; Length 542;

Best Local Similarity 68.2%; Pred. No. 3.4e-145;

Matches 380; Conservative 70; Mismatches 83; Indels 24; Gaps 6;

Qy 1 MKMASNDAPSDNGAAGLVPPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNPFVQAP 60

Db 1 MKMASNDAPSDTGAAGLVPESNEVMALEPVAGAAIAAPVTGQNIIDPWIRANFVQAP 60

Qy 61 GGEFTVSPRNGPGEVLLNLELGPENPYLAHLARMYNGVAGGFVQVVLGNAPTAKTII 120

Db 61 NGEFTVSPRNPAGEVLLNLELGPENPYLAHLARMYNGVAGGMEVQVMLGNAPTAKGLV 120

Qy 121 FAAIPNPFPIDNLSAQITMCPHVIQVROLEPNLPHDPVDRNPFHYNQGSRLRLTA 180

Db 121 FAAVPPHPFVENLSPKQITMFPFHVIDVRLLEPLLPLPFDVRSNPFHYNQKDDPKMIRVA 180

Qy 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTKPFTLPILTISEMSNS 240

Db 181 MLYTPLRANSGDDVFTVSCRVLTRPSDFDTYLVPPTVESKTKPFTLPILTISELSNS 240

Qy 241 RFPVPFIESLHPTSPENIVVQCGNGRVTLDGELMGTTQLLPSCICAFRGVLTRSTGRASDQ 300

Db 241 RFPVPIDQMYTSPNEVISVQCGNGRCTLGDELGGTQLQVSGICAFKGEV---TAHLHDN 297

Qy 301 ADTATPRLFNYYVHVLQNLNGTVPDPAEDIPGLGTDPFRCKVFGVASQR-----NL 353

Db 298 -----DHLNNVTITNLNGSPDPSEDIPAPLGVDPFGQGRVFGVISRDKQNAAGHS 348

Qy 354 DSTTTRAEAKVDTAGRTFKLGSLEIST-DSDDDFDQNPQTKFTVPVIGVDNEAB-FQOM 411

Db 349 EPANRGRHDAVPTTYTAQYTPKLGQIQIGTWQTDLLTVNQPKVFTPVGL---NDTEHFQW 405

QY 412 SLPDYSGQFTHNMNLAPAVAFNPFGEQLLFFRSQLPSSGGRSNGVLDCLVPOEWQHIFYQ 471  
Db 406 VVPRVAGALNTNLAPSVAPVPPGERLLFFRSHLPLKGGYGNAIDCLLPQEWVQHFYQ 465  
QY 472 ESAPQATQVALVRYNPDGTGKVLFFAKLHKLGMFTIANNNGDSPIVPPNGYFRFESWNP 531  
Db 466 EAAPSMSEVALVRYNPDGTGKVLFFAKLHKLGMFTIANNNGDSPIVPPNGYFRFESWNP 525  
QY 532 FYTLAPMGTCNGRRRIQ 548  
Db 526 FYSLAPMGTCNGRRRVQ 542  
RESULT 2  
S40111  
capsid protein - human calicivirus (strain Bristol isolate B493)  
C;Species: human calicivirus  
A;Variety: strain Bristol isolate B493  
C;Date: 25-Dec-1994 #sequence\_revision 27-Feb-1997 #text\_change 28-Jul-2000  
C;Accession: S40111  
R;Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
submitted to the EMBL Data Library, December 1993  
A;Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de  
A;Reference number: S40111  
A;Accession: S40111  
A;Molecule type: genomic RNA  
A;Residues: 1-539 <GRE>  
A;Cross-references: EMBL:X76716; NID:G436410; PIDN:CAAS4134.1; PID:G436411  
A;Experimental source: human enteric calicivirus strain Bristol isolate B493  
C;Superfamily: human calicivirus capsid protein  
C;Keywords: capsid protein; coat protein  
Query Match 67.2%; Score 1947.5; DB 2; Length 539;  
Best Local Similarity 66.1%; Pred. No. 1.le-138;  
Matches 364; Conservative 74; Mismatches 94; Indels 19; Gaps 4;  
QY 1 MKWASNDAAPSNDGAAG---LVPEIN-NEAMALDPVAGAAIAAPLTGQOQNIIDPWIMNNFVQAP 60  
Db 1 MKWASNDANPSDGAAGLVPEINNEAMALDPVAGAAIAAPLTGQOQNIIDPWIMNNFVQAP 60  
QY 61 GGEFTVSPRNSPGEVLLNLELGPENIPYLAHLARMYNGYAGGFEVQVVLGNAFTAGKII 120  
Db 61 GGEFTVSPRNPAGEILWSAPLGPDLNLYLHLSRMVNGYAGGFEVQVVLGNAFTAGKVI 120  
QY 121 FAALPPNPPIDNLGAAQITMCPHVIVDVROLEPNLMPDVRNFFHYNGSDRLRLIA 180  
Db 121 FAAVPPNPPTEGLSPSQVTFPHIIVDVROLEPNLMPDVRNFFHYNGQANDSTLKLIA 180  
QY 181 MLYTPLRANNSGDDVFTVSCRVLTRPSPDFSNFLVPPVTSKTKPFTLPILITISEMSNS 240  
Db 181 MLYTPLRANNGDDVFTVSCRVLTRPSPDFDFILVPPVTSKTKPFTVPLVIEEMNS 240  
QY 241 RFPVPIESLHTSPTEINIVVQCGRVTLDGELMGTTLQSPQICAFRGVLTSTRASDQ 300  
Db 241 RFPVPLEKLYGPSSAFVVPQNGKCTTDGVLGTLQLSAUNIENFRGDVTHIAG----- 295  
QY 301 ADTATPRLFNYYHVLQDNLNGTPYDPAEDIPGLTDPDRGKVGVASQ-RNLDSTTRA 359  
Db 296 -----SDHYTMNLASQNSNYDPTFEEIPAPLGTDPDFVKGKIQGLLTQTTRADGSTRA 346  
QY 360 HEAKVDTTTRAEKAVDTTTRAGRTPKLGSLEISTDSDDDQNPQTKFTPVGIGVDN----- 404  
Db 347 HKATVSTGSVHTFPKLGSVQVQTTDNDNFQAGQNTKFTPVGVIQDGDHQQNEPQWLLPN 406  
QY 416 YSGQFTHNMNLAPAVAFNPFGEQLLFFRSQLPSSGGRSNGVLDCLVPOEWQHIFYQ 475  
Db 407 YSGRTGHNVLAPAVAPTTPGQQLLFFRSTWPGCGSGYPNNMLDCLLPQEWVHLHYQGAAP 466  
QY 476 ACTQVALVRYNPDGTGKVLFFAKLHKLGMFTIANNNGDSPIVPPNGYFRFESWNPFFYL 535  
Db 467 AOSDVALHKLFFNPDGTGKVLFFAKLHKLGMFTIANNNGDSPIVPPNGYFRFESWNPFFYL 526  
QY 536 APMGTGNGRRR 546

Db 527 APMGTGNGRRR 537  
RESULT 3  
B37471  
capsid protein - Norwalk virus  
C;Species: Norwalk virus  
C;Date: 24-Feb-1994 #sequence\_revision 25-Apr-1997 #text\_change 28-Jul-2000  
C;Accession: B37471  
R;Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.  
Virology 195, 51-61, 1993  
A;Title: Sequence and genomic organization of Norwalk virus.  
A;Reference number: A37471; MUID:93303939; PMID:8391187  
A;Accession: B37471  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: genomic RNA  
A;Residues: 1-530 <JIA>  
A;Cross-references: GB:M87661; NID:g1061311; PIDN:AAB50466.1; PID:g1061313  
A;Note: sequence extracted from NCBI backbone (NCBIP:134157)  
C;Superfamily: human calicivirus capsid protein  
Query Match 41.8%; Score 1211; DB 2; Length 530;  
Best Local Similarity 45.4%; Pred. No. 2.9e-83;  
Matches 260; Conservative 74; Mismatches 161; Indels 78; Gaps 11;  
QY 1 MKWASNDAAPSNDGAAG---LVPEIN-NEAMALDPVAGAAIAAPLTGQOQNIIDPWIMNNF 56  
Db 1 MMWASKATSSVDGASGAGQLVPEVNASDPLAMPVAGSSTAVATAGQVNPIDPWIMNNF 60  
QY 57 VQAPGGEFTVSPRNSPGEVLLNLELGPENIPYLAHLARMYNGYAGGFEVQVVLGNAFTA 116  
Db 61 VQAPGGEFTTSPNTPGCVLFDLSLGHLPNPFLLHLSQMYNGWGNVRVIMLAGNAFTA 120  
QY 117 GKIIIFAAIPNFPIDNLGAAQITMCPHVIVDVROLEPNLMPDVRNFFHYNGSDRL 176  
Db 121 GKIIIVSCIPPGGSHNLITIAQATLFFPHIADVTLDIEVPLEDVRNVLPHNDRNQOTM 180  
QY 177 RLIAMLYTLRANNSGDDVFTVSCRVLTRPSPDFSNFLVPPVTSKTKPFTLPILITISE 236  
Db 181 RLVCMLYTLRATGGTGDGDFVAGRVMTCPSPDFNLFVPPVTEQKTRPFTLPNPLSS 240  
QY 237 MSNRFVPIESLHTSPTEINIVVQCGRVTLDGELMGTTLQSPQICAFRGVLTSTR 296  
Db 241 LSNRAPLPISSIGISPDNVQSVQFQNGKCTLDGRLVGTTPVSLSHVAKIRG-----TSN 295  
QY 297 ASDQADTATPRLFNYYHVLQDNLNGTPYDPAEDIPGLTDPDRG-----KVFGVAS 349  
Db 296 GT-----VINLTDLGDTDFHPFEG-PADIGPFDLGGCDWHINMTQFGHSS 339  
QY 350 QRNLSTTTRAEKAVDTTTRAGRTPKLGSLEISTDSDDDQNPQTKFTPVGIGVDN----- 404  
Db 340 QTQYD-----VDITPDTFVPHLGSIQAN-----GIGSGNYGVVL 373  
QY 405 -----EAFQOQSLPDYSGQFTHNMNLAPAVAFNPFGEQLLFFRSQLPSSGGRSN 454  
Db 374 SWISPPSHSGSQVDLWKIPNYGSSITATHLAPSVYPPGFEGLVFFMSXMPGFCAYN- 432  
QY 455 GVLDCLVPOEWQHIFYOESAPATQVQVALVRYNPDGTGKVLFFAKLHKLGMFTIANNNGS- 513  
Db 433 --LPCLLPQEIYSHLASEQAQPTVGEAALLHYVDPDTRNLGFEKAYPDGFLTCVPGNGASS 490  
QY 514 -PIVPPNGYFRFESWNPFFYTLAPMGTCNGRR 545  
Db 491 GPQQLPFGVFFVSVWSRFPYQLKPVGTASSAR 523  
RESULT 4  
B37491  
major capsid protein [similarity] - Southampton virus  
N;Alternate names: orf2 protein  
C;Species: Southampton virus  
C;Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 18-Aug-2000



338 PDGRKGVGVASQNRDLSTTRAHEAKVDVT-----AGRFTPKLGSLEISTDSDDFQDN 390  
454 POTTITAMTASNGVDYTVAEYRTNNGTHPKGFVINGNLTKV-----KGSNDLGET 507  
391 QPTK---FTPV 399  
508 QOTSRTLFASVG 519  
RESULT 8  
S55399  
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)  
C;Species: rabbit hemorrhagic disease virus  
A;Variety: isolate BS89  
C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 18-Jun-1999  
C;Accession: S55399  
R;Rossi, C.  
submitted to the EMBL Data Library, May 1995  
A;Reference number: S55399  
A;Accession: S55399  
A;Molecule type: genomic RNA  
A;Residues: 1-2344 <ROS>  
A;Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641  
A;Experimental source: isolate BS89  
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C;Keywords: polyprotein  
Query Match 9.6%; Score 277.5; DB 2; Length 2344;  
Best Local Similarity 27.0%; Pred. No. 4.6e-12;  
Matches 113; Conservative 57; Mismatches 174; Indels 75; Gaps 20;  
QY 9 APSNDGA----AGLVPEINNEAMALDP-----VAGAAATAPLTGOONI 47  
DB 1773 APQCEAAGTATTASVPGTTTD--GLDFGVVATTSVTAENSSASIAIAGIGGP-PQQVDQ 1829  
QY 48 IDPWIMNFVQAPCGEFTVSPRNSPGEVLLNLELGPINPILAHARMYNGYAGGFVQV 107  
DB 1830 QETWRTNFY---NDVFTWSVADAPGSILYTVQHSPPQNNPFTAVLSQMYAGWAGMQFRF 1886  
QY 108 VLAGNAFTAGKIIFAAIPPNPFDNLNSAAQITMCQPHVIVDVRQLEPNVLPMPDVRRNFFH 167  
DB 1887 IVAGSGVFGRLVAIVPPGIEG--PGLEVRQPHVVIDARSLEPVTITMPDLRPNMYH 1944  
QY 168 YNQGSDSRLRLIAMLITPLRANSGDDVFTVSCVCLTRPSPDRSFNVLPP---TVESKT 224  
DB 1945 PTGDEPLVTLVLSVNNL-INPFGGTSATQVTRPSEDFEFVIRAPSSKTVDIS 2003  
QY 225 KP--FTLPILTISEMSNSRFPVPIESLHTSPENIVVQCGNRVTLDELMTGTTQLLP 282  
DB 2004 PAGLITTPVLT-GVGNDRNWNGQIVGL--QFVPGGFSTC-NRHWNLNGSTYGMSSSPRAD 2059  
QY 283 ICAPRGVLTSTRASDAQDTATPLRFNYVHVQDLNLTGTPYDPAEDIPGLGTPD--- 339  
DB 2060 IDHRRG-----SASYPGSGNATNVL--QFWYA---NAGSAVDNPISQV-APDGPDM 2106  
QY 340 ---FRGK-----VGVASQNRDL-----STTRAHEAKVDTTAGRTFKLGSLEISTDS 384  
DB 2107 VPFNGPGIPAAWGVFGAIWNSGAPNVTTVQAYE-----LGFATGAPGNLQPTTNT 2159  
RESULT 9  
S64740  
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)  
N;Contains: VP60 protein  
C;Species: rabbit hemorrhagic disease virus  
A;Variety: isolate AST/89  
C;Date: 12-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 20-Jun-2000  
C;Accession: S64740; S46944; S49018; S65012  
R;Casals, R.; Martin-Alonso, J.; Boga, J.; Parra, F.  
submitted to the EMBL Data Library, May 1995  
A;Description: Genomic organization of rabbit hemorrhagic disease virus determined by d  
A;Reference number: S64740  
A;Accession: S64740

1829 QETWRTNFY---NDVFTWSVADAPGSILYTVQHSPPQNNPFTAVLSQMYAGWAGMQFR 1885  
107 VVLAGNATAGKIIFAAIPPNPFDNLNSAAQITMCQPHVIVDVRQLEPNVLPMPDVRRNFF 166  
1886 FIVAGSGVFGRLVRAVPPGIEG--PGLEVRQPHVVIDARSLEPVTITMPDLRPNMY 1943  
167 HYNQGSDSRLRLIAMLITPLRANSGDDVFTVSCVCLTRPSPDRSFNVLPP---TVESK 223  
1944 HFTGDPGLVPTLVLSVNNL-INPFGGTSATQVTRPSEDFEFVIRAPSSKTVDIS 2002  
224 TKP--FTLPILTISEMSNSRFPVPIESLHTSPENIVVQCGNRVTLDELMTGTTQLLP 281  
2003 SPAGLLTTPVLT-GVGNDRNWNGQIVGL--QFVPGGFSTC-NRHWNLNGSTYGMSSSPRFG 2058  
282 QICAFRGVLTSTRASDAQDTATPLRFNYVHVQDLNLTGTPYDPAEDIPGLGTPD--- 339  
2059 DIDHRRGSASYSGSNATNVLQ-----GSAIDNPISQVAPDGFDDMS 2105  
340 ---FRGK-----VGVASQNRDL-----STTRAHEAKVDTTAGRTFKLGSLEISTDS 384  
2106 FVPFNGPGIPAAWGVFGAIWNSGAPNVTTVQAYE-----LGFATGAPGNLQPTTNT 2159  
RESULT 7  
A48562  
coat protein - San Miguel sea lion virus (serotype 1)  
N;Alternate names: capsid protein  
C;Species: San Miguel sea lion virus  
C;Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
C;Accession: A48562  
R;Neill, J.D.  
Virus Res. 24, 211-222, 1992  
A;Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel  
A;Reference number: A48562; MUID:92410750; PMID:1529644  
A;Accession: A48562  
A;Molecule type: genomic RNA  
A;Residues: 1-702 <NEI>  
A;Cross-references: GB:M87481; NID:g334882; PIDN:AAAL6217.1; PID:g334884  
A;Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBI:P113565)  
C;Superfamily: feline calicivirus coat protein  
C;Keywords: capsid protein; coat protein; glycoprotein  
F:208,481,493,545/Binding site: carbohydrate (Aan) (covalent) #status predicted  
Query Match 9.7%; Score 281.5; DB 1; Length 702;  
Best Local Similarity 25.9%; Pred. No. 4e-13;  
Matches 112; Conservative 57; Mismatches 160; Indels 103; Gaps 19;  
QY 8 AAPNDGAAGLVPEINNEAMALDPV---AGAAATAAPLTGOONIIDPWIMNFVQAPGGE 63  
DB 151 AESDGPFGADIVTEQGTIVQQQVPQAQSALTTLAAASTGK--TVD-----CE 196  
QY 64 FTV-----SPRNSPGEVLLNLELGPINPILAHARMYNGYAGGFVQVVLGNA 113  
DB 197 WTTFFSYHTAVNWSSTEAQKILSRALSPELNPLRHISLSYTWSSGGIDVRFVSGSG 256  
QY 114 FTAGKIIFAAIPPNPFDNLNSAAQITMCQPHVIVDVRQLEPNVLPMPDVRRNFFHYNQSD 173  
DB 257 VFGKLAALIVPPG--IEFVESPTMLQYPHVLFDAQTEPVITFIDIRKTLHSMDDTD 314  
QY 174 SRLRLIAMLITPL-RANNSGDDVFTVSCVCLTRPSPDRSFNVLPPVPTVESKTPFTLPIL 232  
DB 315 T-TRUIVIVYNNELINPYEQSEPKSSCSITVETRPSDDFTSLKPP--GSLLKHSIP-- 369  
QY 233 TISEMSNSRFPVPIESLH-----TSPTENIVVQCGNRVTLDELMTGTTQLLSQICAF 286  
DB 370 --SDL-----IPRNSRHWNGNRWSTIDGFGVQPR----- 397  
QY 287 RGVLTSTRASDAQDTATPLRFNY--WHVOLDNLN-GTPYDPAED-----IPG-PLGT 337  
DB 398 ----VFQSNRHHDFDSTTTGWSTPTPIPIEVLTLEKLDGGQYFKYTDTEKSLVFGCLPDGW 453



A:Molecule type: genomic RNA  
A:Residues: 1-2344 <CBO>  
A:Cross-references: EMBL:249271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552  
A:Experimental source: isolate AST/89  
R:Boga, J.; Casais, R.; Martin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.  
submitted to the EMBL Data Library, July 1993  
A:Description: Molecular cloning, sequence and expression of the capsid protein gene from  
A:Reference number: S46944  
A:Accession: S46944  
A:Molecule type: genomic RNA  
A:Residues: 1650-2344 <BO>  
A:Cross-references: EMBL:224757; NID:g515622; PIDN:CAA80881.1; PID:g515623  
A:Experimental source: isolate AST/89  
R:Parra, F.; Boga, J.A.; Martin, M.S.; Casais, R.  
Virus Res. 27, 219-228, 1993  
A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp  
A:Reference number: S49018; MUID:93255896; PMID:8488721  
A:Accession: S49018  
A:Molecule type: genomic RNA  
A:Residues: 1650-1796 <PAR>  
A:Cross-references: EMBL:224757  
A:Experimental source: isolate AST/89  
A:Accession: S65012  
A:Molecule type: protein  
A:Residues: 1767-1779; 1875-1877, 'X', 1879-1881, 1936-1938, 'X', 1940-1941 <PAM>  
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C:Keywords: blocked amino end; polyprotein  
Query Match 9.5%; Score 276; DB 2; Length 2344;  
Best Local Similarity 22.8%; Pred. No. 66-12;  
Matches 109; Conservative 60; Mismatches 161; Indels 148; Gaps 19;  
QY 9 APSNDGA---AGLVEINNEAMALDP-----VAGAAIAAFLTCQNI 47  
DB 1773 AQGEAAGTATTASVGTITDGM--DPGVATTSVVTAENSSASIAATAGIGP-PQVDQ 1829  
QY 48 IDPWNNVQAPGGEFTVSPRNSPGEVLLNLLELGPINPYLAHARMYNGVAGFEVQV 107  
DB 1830 QETWRTNFY---NDVFTWSVADAPGSILYTVQHSPPNNPFTAVLSQMYAGWAGGMQFRF 1886  
QY 108 VLAGNAFTAGKIFRAIPNFPIDNLSAAQITMCCHVIVDVROLEPNLMPDVRNPF 167  
DB 1887 IVAGIGVFGRLVAIVPGIEIG--PGLEVQFPFHVIDARSLPVTITMPLDLPNMYH 1944  
QY 168 YNQGSRLRLIAMLTYPLRANNSGDVFTVSCRVLTRPSPPFSFNFLVPP---TVESKT 224  
DB 1945 PTGDCPLVTLVSVNNL-INPFGGSTSAIQVTVETRESEDEFWMIRAPSKTVDSLS 2003  
QY 225 KP--FTLPILTISEMSNSRFPVPIESLHTSPTENIVVOCONGRVTLDBELMGTQLLPSQ 282  
DB 2004 PAGLLTTPVLT-GVGNDNRW-----NGQIVG-LQVPVG- 2034  
QY 283 ICAFRVLTSTRASDAQDTATPRLFNYVHVDNLNGTYPDPAEDIPGLTPTDFRG 342  
DB 2035 -----GFST-----CNRHW-----NLNGSTY----- 2050  
QY 343 KVFVGVAQRNLSDTTTRAHEAKVDTTAGRTPTPKLSLEISTDSDDFDQNPQTK-----FTP 397  
DB 2051 -----GWSPPFGDIGHRRGASYPGNATVNLQFWYAN 2084  
QY 398 VGIGVDNE-AEFQWQSLPPYSGQFTHNMNLAPAVAPNPFEGQLLFFRSQLPSSGGRSN 454  
DB 2085 AGSAIDNPISQVAPDGFDP-----MSFVPFNGPGIPAAAGWVGFGAIWNSGAPN 2134  
RESULT 10  
VCWMF9  
coat protein - feline calicivirus (strain F9)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 28-Jul-2000  
C:Accession: B43382; C45538; PQ0407; S23702  
R:Cartier, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.

Virology 190, 443-448, 1992  
A:Title: The complete nucleotide sequence of a feline calicivirus.  
A:Reference number: A43382; MUID:92410623; PMID:1529544  
A:Accession: B43382  
A:Molecule type: genomic RNA  
A:Residues: 1-671 <CAR1>  
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
R:Cartier, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.  
Arch. Virol. 122, 223-235, 1992  
A:Title: Identification and sequence determination of the capsid protein gene of feline  
A:Reference number: A45538; MUID:92117861; PMID:1731695  
A:Accession: C45538  
A:Molecule type: genomic RNA, protein  
A:Residues: 1-671 <CAR2>  
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
A:Experimental source: strain F9  
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)  
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.  
J. Gen. Virol. 73, 2429-2433, 1992  
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel  
A:Reference number: PQ0407; MUID:93019069; PMID:1402818  
A:Accession: PQ0407  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>  
A:Cross-references: PIDN:AAB23553.1; PID:g257083  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 9.5%; Score 275.5; DB 1; Length 671;  
Best Local Similarity 22.3%; Pred. No. 1.1e-12;  
Matches 123; Conservative 70; Mismatches 174; Indels 185; Gaps 23;  
QY 67 SPRNSPGEVLLNLELGPINPYLAHARMYNGVAGFEVQVVLGAGNAFTAGKIIPAAIIP 126  
DB 179 SPSTQGGKILFKQSLGSLPLNLYLEHLAKLYVAMSGSIEVRFSISGSGVFGKLAIVVPP 238  
QY 127 NPPIDNLSAAQITMCCHVIVDVROLEPNLMPDVRNPFNYNQGSRLRLIAMLTYPL 186  
DB 239 G--VDPVQSTSMLOYPHVLFDARQVEPVIFCLPDLRLSTLYHLMSDITD-TSLVIMVYNDL 295  
QY 187 RANNSGDDVFTVSC--RVLTSPSPFSFNFLVPP-----TVES-----KT----- 224  
DB 296 -INPYANDANSNGCIVTVETKPGDPKFKHLLKPPGSMULTHGSIPSDLPKTSLSLWIGNRY 354  
QY 225 -----KPRTL-----PI-LTISEMSNSRFPVPIESLHTSP 253  
DB 355 WSDITDFVIRPEVFOANRHFDFNQETAGWSTPRFRPISVTITEQNGAKLIGV-----A 408  
QY 254 TENIVVQCQNG--RVTLDGELMGTQLLPSQICAPRGVLTSTRSTRASDAQDTATPRLFNY 311  
DB 409 TDYIVPGIDPQWPDITIGEL-----IPAGDYAITNCTGN----- 443  
QY 312 YMHVQDLNLTGTPYDPAEDIPGLTPTDFRGKVGKVGASQRNLSDTTTRAHEAKVDTTAGRF 371  
DB 444 -----DITTATGYDTADIIK--NNTNFRGMVYICGSLQRAMG-----DKKISNTAFIT 488  
QY 372 TPKLSLEISTDSDP-----PDQNPQTKFTPVGIGVDNEAEFQWQSLPDYSGQFTH 422  
DB 489 TATL-----DGDNNKINPCNTDQSKIVVFDNHHVKGKAQTSDDTLALLGYTG----- 537  
QY 423 NNNLAPAVAPNPFPGQLL-----FFRSQLPSSGGRSNGLVDCLVPQEWQVH--FYQES 473  
DB 538 -----IGEQAIGSDRDRVVRISTLEPCTGARGG-----NHPIFYKNS 573  
QY 474 AP-----AQTQVALVRYVNP-----DTGKVLFEAKLHLKFWMT 506  
DB 574 IKLGVVIRSIDVFNSEQLLHTSRQLSLNHYLLPPDSFAVVRIIDNSGWNFDIGSDSGFSF 633  
QY 507 IANNQDSITVP 518  
DB 634 VGVSGFGKLEFP 645

C;Keywords: capsid protein; coat protein; glycoprotein  
F;177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.0%; Score 262; DB 1; Length 668;  
Best Local Similarity 34.4%; Pred. No. 1.1e-11;  
Matches 54; Conservative 32; Mismatches 59; Indels 12; Gaps 4;

Qy 67 SPNSGCEVLLNLELGPENIPYLAHLARMYNGYAGGFEVQVVLVAGNAFTAGKIIFAAIPP 126  
Db 179 STSETQKILFKQSLGFLNPNLYTHLAKLYVAMSGVDVRFSSGSGVGGKLAIVVPP 238  
Qy 127 NFPIIDNLAAQITMCPHVIVDVRQLSPVNLPMQVNRNFFHYNQSGDSRLRLIAMLYTPL 186  
Db 239 G--IDPVQSTSMQLQYPHVLFDAQVPEVIFSDLRSTLYHLMSTDTT-TSLVIMVNDL 295  
Qy 187 -----RANSGDDVFTVSCVRLTRPSPDSEFNFLVPP 218  
Db 296 INPYANDSNSSGCIIVTVE----TKPGDPFKFHLKPP 328

RESULT 13  
VCWNFC  
coat protein - feline calicivirus (strain Japanese F4)  
N;Alternate names: capsid protein  
C;Species: feline calicivirus  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C;Accession: B40481  
R;Tohya, Y.; Taniuchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamaza  
Virology 183, 810-814, 1991  
A;Title: Sequence analysis of the 3'-end of feline calicivirus genome.  
A;Reference number: A40481; MUID:91306470; PMID:11853578  
A;Accession: B40481  
A;Molecule type: genomic RNA  
A;Residues: 1-668 <TOH>  
A;Cross-references: GB:D50357; NID:g221264; PIDN:BAA14371.1; PID:g221266  
C;Superfamily: feline calicivirus coat protein  
C;Keywords: capsid protein; coat protein; glycoprotein  
F;177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.0%; Score 261.5; DB 1; Length 668;  
Best Local Similarity 30.3%; Pred. No. 1.2e-11;  
Matches 59; Conservative 41; Mismatches 78; Indels 17; Gaps 6;

Qy 30 DPVAGAAIAAPLTGQONTIDPW-INNVFVQAPGGEFTVSPNSGCEVLLNLELGPENPY 88  
Db 145 EPSAQWSTAADWASGKSVDSWEAFFSPTS----VNMSTSETQKILFKQSLGFLNPNY 200  
Qy 89 LAHLARMYNGYAGGFEVQVVLVAGNAFTAGKIIFAAIPPFPIDNLAAQITMCPHVIVDV 148  
Db 201 LEHLAKLYVAMSGSIEVRFSSGSGVFGKLAIVVPPG--VDPVQSTSMQLQYPHVLFDA 258  
Qy 149 ROLPEVNLPMQVNRNFFHYNQSGDSRLRLIAMLYTPL-----RANSGDDVFTVSCVRL 203  
Db 259 RQVEPVIFIPDLRSTLYHVMSTDTT-TSLVIMVNDLINPYANDSNSSGCIIVTVE---- 313  
Qy 204 TRPSDPDFSNFLVPP 218  
Db 314 TKPGDPFKFHLKPP 328

RESULT 14  
JQ2354  
capsid protein - feline calicivirus (strain NADC)  
C;Species: feline calicivirus  
C;Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
C;Accession: JQ2354  
R;Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.  
J. Gen. Virol. 74, 2519-2524, 1993  
A;Title: Analysis of feline calicivirus capsid protein genes: identification of variable  
A;Reference number: JQ2354; MUID:94065683; PMID:7504075  
A;Accession: JQ2354  
A;Molecule type: mRNA  
A;Residues: 1-668 <SEA>

RESULT 11  
C48562  
coat protein - San Miguel sea lion virus (serotype 4)  
N;Alternate names: capsid protein  
C;Species: San Miguel sea lion virus  
C;Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
C;Accession: C48562  
R;Neill, J.D.  
Virus Res. 24, 211-222, 1992  
A;Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea  
eins.  
A;Reference number: A48562; MUID:92410750; PMID:1529644  
A;Accession: C48562  
A;Molecule type: genomic RNA  
A;Residues: 1-703 <NEI>  
A;Cross-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888  
A;Note: sequence extracted from NCBI backbone (NCBIP:1113567)  
C;Superfamily: feline calicivirus coat protein  
C;Keywords: capsid protein; coat protein; glycoprotein  
F;89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.3%; Score 269; DB 1; Length 703;  
Best Local Similarity 32.3%; Pred. No. 3.5e-12; Indels 18; Gaps 7;  
Matches 71; Conservative 38; Mismatches 89;

Qy 8 AAPSNDGAAGLVE-----INNEAMALDPVAGAAIAAPLTGQONTIDPWNNFVQAPGGE 63  
Db 151 AESDGFSGSAIEVTEEGTVVQQQPAPAPTALATATAGTG-KSVEQEWMTFFSYHTSIW 209  
Qy 64 FTVSPNSGCEVLLNLELGPENIPYLAHLARMYNGYAGGFEVQVVLVAGNAFTAGKIIFAA 123  
Db 210 STV---ESQKILYQALNPSINPYLDHIAKLYSTWSGGIDVRFTVSGSGVFGKLAALL 266  
Qy 124 IPNPF-PIDNLAAQITMCPHVIVDVRQLSPVNLPMQVNRNFFHYNQSGDSRLRLIAML 182  
Db 267 VPPGVEPIESVSMQLQY---PHVLFDAKQTEPVIFTPIDIRKTLFHSMDTDT-TKLVINP 322  
Qy 183 YTPLRANSGDDVFTVSCVRLTRPSPDSEFNFLVPP 218  
Db 323 Y-----ENGVENKTTCSITVETRFSADETFALLKPP 353

RESULT 12  
VCWNFF  
coat protein - feline calicivirus (strain CFI/68 FIV)  
N;Alternate names: capsid protein  
C;Species: feline calicivirus  
C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
C;Accession: A40507; B40507; T09246  
R;Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.  
J. Virol. 65, 5440-5447, 1991  
A;Title: Nucleotide sequence and expression of the capsid protein gene of feline caliciv  
A;Reference number: A40507; MUID:91374597; PMID:1716692  
A;Accession: A40507  
A;Molecule type: genomic RNA  
A;Residues: 1-668 <NEI>  
A;Cross-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875  
A;Accession: B40507  
A;Molecule type: protein  
A;Residues: 373-379;403-419;481-489;560-566 <NE2>  
R;Neill, J.D.  
submitted to the EMBL Data Library, April 1998  
A;Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.  
A;Reference number: Z16626  
A;Accession: T09246  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: genomic RNA  
A;Residues: 1-668 <NE3>  
A;Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAAL3993.1; PID:g537256  
A;Experimental source: strain CFI/68 FIV  
C;Superfamily: feline calicivirus coat protein

A:Cross-references: GB:L09718; NID:G305104; PIDN:AAA16485.1; PID:G305105  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein

```
Query Match          8.9%; Score 257; DB 2; Length 668;
Best Local Similarity 23.6%; Pred. No. 2.6e-11;
Matches 72; Conservative 52; Mismatches 99; Indels 82; Gaps 10;

QY 67 SPNRSPEVLLNLELGPEINPYLAHLMARYNGYAGGFEVQVVLGAGNAFTAGKIIFAAIIPP 126
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 179 STSETQKILFKQSLGPLENPLYELHKLKLYVAWSGSVEVRFSGSGVFGGKLAIVVPP 238
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 127 NFPIDNLSAAQITMCHVIVDVRQLEPNLMPDVRRNFFHYNQSGDSRLRLIAMLYTEL 186
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 239 G--VDPVQSTSLMQYPHVLFDAQVDFVIFSPDLRSTLYHLMPDQTD-TSLVIMVYNDL 295
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 187 -----RANNSGDDVFTVSCRVLTRPSPDFSFNPLVPP----- 218
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 296 INPYANDSNSGCIVTVE---TKGPDFKFHLLKPPGSMLTGSGVSPDLIPKSSSLWIG 351

QY 219 -----TVESKTKPFTL-----PI-LTISEMSNSRFPVPFIESLH 250
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 352 NRYWSDITDFVVRPFVQANRHFNQETAGNSAPRFRPITITISESKGSLGIGVATDY 411

QY 251 -----TSPTENIVVQCONGRVTLGELMGTTQLLPSCAIFRGV-LTR 292
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 412 IVGIPDGWPDTTIAEDLTAPAGDYAITSGNGNDITTGSEYDSTEVIKNN-TNFRGMYICG 470

QY 293 STSRA 297
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 471 SLQRA 475
```

RESULT 15  
JQ2356  
capsid protein - feline calicivirus (strain KCD)  
C:Species: feline calicivirus  
C>Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
C:Accession: JQ2356  
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.  
J: Gen. Virol. 74, 2519-2524, 1993  
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable  
A:Reference number: JQ2354; MUID:94065683; PMID:7504075  
A:Accession: JQ2356  
A:Molecule type: mRNA  
A:Residues: 1-668 <SEA>  
A:Cross-references: GB:L09719; NID:G305107; PIDN:AAA16487.1; PID:G305108  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein

```
Query Match          8.8%; Score 255; DB 2; Length 668;
Best Local Similarity 33.8%; Pred. No. 3.6e-11;
Matches 53; Conservative 31; Mismatches 61; Indels 12; Gaps 4;

QY 67 SPNRSPEVLLNLELGPEINPYLAHLMARYNGYAGGFEVQVVLGAGNAFTAGKIIFAAIIPP 126
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 179 STSETQKILFKQSLGPLENPLYELHKLKLYVAWSGSIEVRFSGSGVFGGKLAIVVPP 238
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 127 NFPIDNLSAAQITMCHVIVDVRQLEPNLMPDVRRNFFHYNQSGDSRLRLIAMLYTEL 186
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 239 G--VDPVQSTSLMQYPHVLFDAQVDFVIFSPDLRSTLYHLMSDQTD-TSLVVMAYNDL 295
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :

QY 187 -----RANNSGDDVFTVSCRVLTRPSPDFSFNPLVPP 218
   |  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :  :
Db 296 INPYANDSNSGCIVTVE---TKGSDFRFHLKPP 328
```

Search completed: June 1, 2004, 13:55:25  
Job time : 12.5738 secs

**THIS PAGE BLANK (USPTO)**



D	b
61	GGEFTVS PRNAPGELLWSAPLGPDLNPYLSHLRMTNGYAGGFVEQVVLGNFATAGKI 120 
121	F AAVPPNFPT EGLSPSQVTMEPHIIIVDROLPEVLIPLDPVRNNFYHYNQSDSTIKLIA 180 
121	F AAVPPNFPT EGLSPSQVTMEPHIIIVDRLQEPVLLIPLDPVRNNFYHYNQSDSTIKLIA 180 
181	M LYTFLRANNA GDDVF TSCRVLT RPSDFDI ELVPPTVESRTK PTPVILTVEMS NS 240 
181	M LYTFLRANNA GDDVF TSCRVLT RPSDFDI ELVPPTVESRTK PTPVILTVEMS NS 240 
241	R PPIPLEKLYT GPSSAFVVQPONGRCITDG VLGCTTGTSOLSAVNICTFRGDVTHI AGSHDYT 300 
241	R PPIPLEKLYT GPSSAFVVQPONGRCITDG VLGCTTGTSOLSAVNICTFRGDVTHI AGSHDYT 300 
301	M NLASQNMSNY DPTBEIPALPGTD PFVGKI QGM LTQTTR EDGSTRA HKATV STGS VHFTP 360 
301	M NLASQNMSNY DPTBEIPALPGTD PFVGKI QGM LTQTTR EDGSTRA HKATV STGS VHFTP 360 
361	K LGSGVOYTT DNND FQTGONT KFTPVGV I QGNHNHONPQQWLP NYSGR TGHNHLAPA 420 
361	K LGSGVOYTT DNND FQTGONT KFTPVGV I QGNHNHONPQQWLP NYSGR TGHNHLAPA 420 
421	V APTEEGEQLLFPR STMPCSGYPNMNLDC LLPOEWQHFCQE AAPAQSDVALLRF VNPD 480 
421	V APTEEGEQLLFPR STMPCSGYPNMNLDC LLPOEWQHFCQE AAPAQSDVALLRF VNPD 480 
481	T GRVLFECK LHKS GGYVTVAHTG PHDLVIPPNYG FRFD SWNVNQ FYTLPM GNAGRRRAL 539 
481	T GRVLFECK LHKS GGYVTVAHTG PHDLVIPPNYG FRFD SWNVNQ FYTLPM GNAGRRRAL 539 
<b>RESULT 2</b>	
Q9W9T8	PRELIMINARY; PRT; 539 AA.
ID AC Q9W9T8	
DT DT 01-NOV-1999	(TrEMBLrel. 12, Created)
DT DT 01-NOV-1999	(TrEMBurel. 12, Last sequence update)
DT DT 01-OCT-2003	(TrEMBurel. 25, Last annotation update)
DE DE Capsid protein.	
OS OS Human calicivirus NLV/Mora/97/SE,	
OS OS Human calicivirus isolates, and	
OS OS Human calicivirus NLV/Oberschleissheim/112/99/DE.	
OC OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;	
OC OC Norovirus	
OX NCBI_TaxID=189467, 150080, 173925;	
[1]	
RP SEQUENCE FROM N.A.	
RN SPECIES=Human calicivirus NLV/Mora/97/SE; STRAIN=Hu/NLV/Mora/97/SE;	
RC Nilesen M., Torven M., Thorhaugen M., Hedlund K.-O., Svensson L.;	
RA "Expression, self assembly and intracellular location in mammalian	
RT cells of a lardsdale virus-like calicivirus protein.";	
RL Submitted (Feb-2002) to the EMBL/GenBank/DDBJ databases.	
[2]	
RP SEQUENCE FROM N.A.	
RN SPECIES=Human calicivirus isolates; STRAIN=Various strains;	
RX MEDLINE=99246332; PubMed=10228052;	
RA Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;	
RT "Identification of a distinct common strain of 'Norwalk-like viruses'	
RT having a global distribution.";	
J. Infect. Dis. 179:1334-1344(1999).	
[3]	
RP SEQUENCE FROM N.A.	
RC SPECIES=Human calicivirus NLV/Oberschleissheim/112/99/DE;	
RA STRAIN=NLV/Oberschleissheim/112/99/DE;	
RC Kuengel U., Hoehne M., Schreiber B.;	
RT "Molecular epidemiology of outbreaks of gastroenteritis associated	
RT with Norwalk-like viruses in Germany.";	
Submitted (Oct-2001) to the EMBL/GenBank/DDBJ databases.	
REML AY081134; AAL87461.1;- -	
DR EM BL; AF080559; AAD40498.1;- -	
DR EM BL; AF080549; AAD40488.1;- -	
DR EM BL; AF080551; AAD40490.1;- -	

Query Match	99.5%;	Score 2872;	DB 12;	Length 539;
Best Local Similarity	99.6%;	Pred. No. 4.7e-215;		
Matches 537;	Conservative	1;	Mismatches	1; Indels 0; Gaps 0;
EMBL; AF080555; AAD40494.1; --				
DR EMBL; AF080557; AAD40496.1; --				
DR EMBL; AF427113; AAL18869.1; --				
DR InterPro; IPR004005; Calici_coat.				
DR InterPro; IPR008975; Viral_cap_coat.				
DR Pfam; PF00915; Calici_coat; 1.				
SQ SEQUENCE 539 AA; 58902 MW; 0C6CDEFA3D6CC345B CRC64;				
QY 1 MKMASNDANPSDGSSTANLVPENVEVNEVMALPEVVGAAIAAIVAGQONVIDPWIRNFFVQAP 60				
DB 1 MKMASNDANPSDGSSTANLVPENVEVNEVMALPEVVGAAIAAIVAGQONVIDPWIRNFFVQAP 60				
QY 61 GGEFTVSPNAPGEILWSAPLGDPLNPLYSLHARMYNGYAGGPEVQVILAGNAFTAGKII 120				
DB 61 GGEFTVSPNAPGEILWSAPLGDPLNPLYSLHARMYNGYAGGPEVQVILAGNAFTAGKII 120				
QY 121 FAAPVPNPFTEGLSPSQVTWTFPHIIVDVRLQLEPVLIIPLDVRRNFYHYNQSNDSSTIKLIA 180				
DB 121 FAAPVPNPFTEGLSPSQVTWTFPHIIVDVRLQLEPVLIIPLDVRRNFYHYNQSNDSSTIKLIA 180				
QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFIPLVPTVESRTKPTVPLTLTVEEMSNS 240				
DB 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFIPLVPTVESRTKPTVPLTLTVEEMSNS 240				
QY 241 RFPPLKLYITGSSAFVQVQNGRCRTDGVLLGTTLQLSAVNICTFRGDVTHIAGSHDYT 300				
DB 241 RFPPLKLYITGSSAFVQVQNGRCRTDGVLLGTTLQLSAVNICTFRGDVTHIAGSHDYT 300				
QY 301 MNLASQWNSNYDPTTEIIPAPLGTFPVGKIQGLTQTTRPDGSGTRAHKATVSTGSHVFTP 360				
DB 301 MNLASQWNSNYDPTTEIIPAPLGTFPVGKIQGLTQTTRPDGSGTRAHKATVSTGSHVFTP 360				
QY 361 KLGSVQYITDNDPQTQNTKFTPVGVIQGNHONPEQQWLPNYSGRGTGHNVHLAPA 420				
DB 361 KLGSVQYITDNDPQTQNTKFTPVGVIQGNHONPEQQWLPNYSGRGTGHNVHLAPA 420				
QY 421 VAPTEPGQLLFFRSTMPGCGYPNNMLDCLLPQEWVQHFCQEAAPASQDVALLRFVNPD 480				
DB 421 VAPTEPGQLLFFRSTMPGCGYPNNMLDCLLPQEWVQHFCQEAAPASQDVALLRFVNPD 480				
QY 481 TGRVLFECKLHKSGYVTVHAHTGPHDLVLPNGYFRFDSWVNOFYTLAPMNGCAGRRAL 539				
DB 481 TGRVLFECKLHKSGYVTVHAHTGPHDLVLPNGYFRFDSWVNOFYTLAPMNGCAGRRAL 539				
RESULT 3				
Q915C5	PRELIMINARY;	PRT;	539 AA.	
ID Q915C5				
AC Q915C5;				
DT 01-DEC-2001 (TrEMBLrel. 19, Created)				
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)				
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)				
DE Capsid protein.				
OS Human calicivirus NLV/Beeskow/124/00/DE.				
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;				
OC Norovirus.				
OX NCBI_TaxID=173931;				
RN [1]				
RP SEQUENCE FROM N.A.				
RC STRAIN=NLV/Beeskow/124/00/DE;				
RA Kuenkel U., Hoshne M., Schreier E.;				
RT "Molecular epidemiology of outbreaks of gastroenteritis associated				
RT with Norwalk-like viruses in Germany."				
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.				
DR EMBL; AF427120; AAL18876.1; --				
DR InterPro; IPR004005; Calici_coat.				
DR InterPro; IPR008975; Viral_cap_coat.				
DR Pfam; PF00915; Calici_coat; 1.				
SQ SEQUENCE 539 AA; 58916 MW; CF6B06EBE45C17D8F CRC64;				

```
Query Match 99.5%; Score 2871; DB 12; Length 539;
Best Local Similarity 99.4%; Pred. No. 5.7e-215;
Matches 536; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPEVNVNMALEPVVGAIAAPVAGQNNIDPWIRNPFVQAP 60
DB 1 MKMASNDANPSDGSSTANLVPEVNVNMALEPVVGAIAAPVAGQNNIDPWIRNPFVQAP 60
QY 61 GGEFTVSPRNAPGEILWSAPLGDPLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNAPGEILWSAPLGDPLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
QY 121 FFAVPPNPFTEGLSPQVTFMFIIVDVROLEPVLIPLDVVRNPFYHYNQSNDSIKLIA 180
DB 121 FFAVPPNPFTEGLSPQVTFMFIIVDVROLEPVLIPLDVVRNPFYHYNQSNDSIKLIA 180
QY 181 MLYTTLRANAGDDVFTVSCRVLTRPSDFDFILVPPVTESRTKPFVPIILTVEMSN 240
DB 181 MLYTTLRANAGDDVFTVSCRVLTRPSDFDFILVPPVTESRTKPFVPIILTVEMSN 240
QY 241 RPPPLEKLYTGPSSAFVQPNQGRCTTGVLLGTTQLSAVNICFRGDTVTHIAGSHDYT 300
DB 241 RPPPLEKLYTGPSSAFVQPNQGRCTTGVLLGTTQLSAVNICFRGDTVTHIAGSHDYT 300
QY 301 MNLASQNNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVFT 360
DB 301 MNLASQNNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVFT 360
QY 361 KLGVSQVYTTDNDFTQNTKFTPVGVIQDGNHNEPQWVLPNYSRGTGHNHVLAPA 420
DB 361 KLGVSQVYTTDNDFTQNTKFTPVGVIQDGNHNEPQWVLPNYSRGTGHNHVLAPA 420
QY 421 VAPTPPGEQLLFFRSTMPGCSGYPNNLDCLLPQEWVQHFCEAAAPQSDVALLFVNP 480
DB 421 VAPTPPGEQLLFFRSTMPGCSGYPNNLDCLLPQEWVQHFCEAAAPQSDVALLFVNP 480
QY 481 TGRVLFECKLHKSGYVTVHTGPHDLVIPPNGYFRFDSVWVQFYTLAPMNGAGRRAL 539
DB 481 TGRVLFECKLHKSGYVTVHTGPHDLVIPPNGYFRFDSVWVQFYTLAPMNGAGRRAL 539

RESULT 4
QYQY55 ID Q8QY55 PRELIMINARY; PRT; 539 AA.
AC Q8QY55;
DT 01-JUN-2002 (TremBLrel. 21, Created)
DT 01-JUN-2002 (TremBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/DJON171/96.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=186843;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/DJON171/96;
RA Nicollier-Jamot B., Pico V., Pothier P., Kohli B.;
RT "Molecular cloning, expression, self-assembly and seroepidemiology of a genogroup II Norwalk-like virus in France.";
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF472623; AAL79839.1;
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58916 MW; 801D18BFA315CAF4 CRC64;

Query Match 99.5%; Score 2871; DB 12; Length 539;
Best Local Similarity 99.4%; Pred. No. 5.7e-215;
Matches 536; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPEVNVNMALEPVVGAIAAPVAGQNNIDPWIRNPFVQAP 60
DB 1 MKMASNDANPSDGSSTANLVPEVNVNMALEPVVGAIAAPVAGQNNIDPWIRNPFVQAP 60
QY 61 GGEFTVSPRNAPGEILWSAPLGDPLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNAPGEILWSAPLGDPLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
QY 121 FFAVPPNPFTEGLSPQVTFMFIIVDVROLEPVLIPLDVVRNPFYHYNQSNDSIKLIA 180
DB 121 FFAVPPNPFTEGLSPQVTFMFIIVDVROLEPVLIPLDVVRNPFYHYNQSNDSIKLIA 180
QY 181 MLYTTLRANAGDDVFTVSCRVLTRPSDFDFILVPPVTESRTKPFVPIILTVEMSN 240
DB 181 MLYTTLRANAGDDVFTVSCRVLTRPSDFDFILVPPVTESRTKPFVPIILTVEMSN 240
QY 241 RPPPLEKLYTGPSSAFVQPNQGRCTTGVLLGTTQLSAVNICFRGDTVTHIAGSHDYT 300
DB 241 RPPPLEKLYTGPSSAFVQPNQGRCTTGVLLGTTQLSAVNICFRGDTVTHIAGSHDYT 300
QY 301 MNLASQNNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVFT 360
DB 301 MNLASQNNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVFT 360
QY 361 KLGVSQVYTTDNDFTQNTKFTPVGVIQDGNHNEPQWVLPNYSRGTGHNHVLAPA 420
DB 361 KLGVSQVYTTDNDFTQNTKFTPVGVIQDGNHNEPQWVLPNYSRGTGHNHVLAPA 420
QY 421 VAPTPPGEQLLFFRSTMPGCSGYPNNLDCLLPQEWVQHFCEAAAPQSDVALLFVNP 480
DB 421 VAPTPPGEQLLFFRSTMPGCSGYPNNLDCLLPQEWVQHFCEAAAPQSDVALLFVNP 480
QY 481 TGRVLFECKLHKSGYVTVHTGPHDLVIPPNGYFRFDSVWVQFYTLAPMNGAGRRAL 539
DB 481 TGRVLFECKLHKSGYVTVHTGPHDLVIPPNGYFRFDSVWVQFYTLAPMNGAGRRAL 539

RESULT 5
QYQY55 ID Q9WR20 PRELIMINARY; PRT; 539 AA.
AC Q9WR20;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus isolates.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=150080;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=358/96015107/1996/FL;
RX MEDLINE=9246332; PubMed=10228052;
RA Noel J.F., Fankhauser R., Liorio-Houde M., Monto A.S., Glass R.I.;
RT "Identification of a novel and distinct common gastroenteritis-like virus, having a globular capsid structure."
RL U.S. Infect. Dis. 273:3339-3343 (1999).
DR EMBL; AF080552; AAD40491.1;
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58875 MW; 0C6CDA3CA1DF95B CRC64;

Query Match 99.3%; Score 2867; DB 12; Length 539;
Best Local Similarity 99.4%; Pred. No. 1.2e-214;
Matches 536; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPEVNVNMALEPVVGAIAAPVAGQNNIDPWIRNPFVQAP 60
DB 1 MKMASNDANPSDGSSTANLVPEVNVNMALEPVVGAIAAPVAGQNNIDPWIRNPFVQAP 60
QY 61 GGEFTVSPRNAPGEILWSAPLGDPLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNAPGEILWSAPLGDPLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120
```

```
QY 121 FAAVPPNPTTEGLSPSQVTMPFHIIIDVRQLEPVLIPLDVVRNNFYHYNQSDSTIKLIA 180
Db 121 FAAVPPNPTTEGLSPSQVTMPFHIIIDVRQLEPVLIPLDVVRNNFYHYNQSDSTIKLIA 180
QY 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDDFDIFLVPPTVESRTKPTFTVPIITVEEMSN 240
Db 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDDFDIFLVPPTVESRTKPTFTVPIITVEEMSN 240
QY 241 RFPPIPLEKLYTGPSAFVVPQNGRCTTGGVLLGTTQLSAVNICTRFGDVTTHIAGSHDYT 300
Db 241 RFPPIPLEKLYTGPSAFVVPQNGRCTTGGVLLGTTQLSAVNICTRFGDVTTHIAGSHDYT 300
QY 301 MNLASQWNSNYDPTTEIPAPLGTDPFVGKIQQMLTQTTREDGSTRAHKATVSTGSHVFTP 360
Db 301 MNLASQWNSNYDPTTEIPAPLGTDPFVGKIQQMLTQTTREDGSTRAHKATVSTGSHVFTP 360
QY 361 KLGSVQYTTDTNNDPOTGQNTKFTPGVITQDGNHQNPEQQWVLPNYSRGTGHNVLHAPA 420
Db 361 KLGSVQYTTDTNNDPOTGQNTKFTPGVITQDGNHQNPEQQWVLPNYSRGTGHNVLHAPA 420
QY 421 VAPTFGEQLLFRSTMPGCGYPNNMLDCLLPQEWVQHFQCEAAQAQSDVALLRFVNDP 480
Db 421 VAPTFGEQLLFRSTMPGCGYPNNMLDCLLPQEWVQHFQCEAAQAQSDVALLRFVNDP 480
QY 481 TGRVLFECKLHKSQYVTVTAHTGPHDLVIPPNNGYFRFDSWVNFQYTLAPMNGAGRRAL 539
Db 481 TGRVLFECKLHKSQYVTVTAHTGPHDLVIPPNNGYFRFDSWVNFQYTLAPMNGAGRRAL 539
QY 539
Db 539
RESULT 6
Q9W9Y9
ID Q9W9Y9 PRELIMINARY; PRT; 539 AA.
AC Q9W9Y9;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus isolates.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_taxid=150080;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=379/96019984/1996/AZ, and 364/96019537/1996/AZ;
RX MEDLINE=93246332; PubMed=10228052;
RA Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;
RT "Identification of a distinct common strain of 'Norwalk-like viruses' having a global distribution."
RL J. Infect. Dis. 179:1334-1344 (1999).
DR EMBL; AF080553; AAD40492.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58872 MW; 1D6CD304101BEF2B CRC64;
Query Match 99.3%; Score 2867; DB 12; Length 539;
Best Local Similarity 99.4%; Pred. No. 1.2e-214;
Matches 536; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
QY 1 MKMASNDANPSDGSSTANLVPEVNEVMALPEVVGAAIAAPVAGQONVIDPWRNNFVOAP 60
Db 1 MKMASNDANPSDGSSTANLVPEVNEVMALPEVVGAAIAAPVAGQONVIDPWRNNFVOAP 60
QY 61 GGEFTVSPRNAGEILWSAPLGPDLNPLYSLHARMYNGYAGFEQVILAGNAFTAGKII 120
Db 61 GGEFTVSPRNAGEILWSAPLGPDLNPLYSLHARMYNGYAGFEQVILAGNAFTAGKII 120
QY 121 FAAVPPNPTTEGLSPSQVTMPFHIIIDVRQLEPVLIPLDVVRNNFYHYNQSDSTIKLIA 180
Db 121 FAAVPPNPTTEGLSPSQVTMPFHIIIDVRQLEPVLIPLDVVRNNFYHYNQSDSTIKLIA 180
QY 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDDFDIFLVPPTVESRTKPTFTVPIITVEEMSN 240
Db 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDDFDIFLVPPTVESRTKPTFTVPIITVEEMSN 240
```

```
Db 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDDFDIFLVPPTVESRTKPTFTVPIITVEEMSN 240
QY 241 RFPPIPLEKLYTGPSAFVVPQNGRCTTGGVLLGTTQLSAVNICTRFGDVTTHIAGSHDYT 300
Db 241 RFPPIPLEKLYTGPSAFVVPQNGRCTTGGVLLGTTQLSAVNICTRFGDVTTHIAGSHDYT 300
QY 301 MNLASQWNSNYDPTTEIPAPLGTDPFVGKIQQMLTQTTREDGSTRAHKATVSTGSHVFTP 360
Db 301 MNLASQWNSNYDPTTEIPAPLGTDPFVGKIQQMLTQTTREDGSTRAHKATVSTGSHVFTP 360
QY 361 KLGSVQYTTDTNNDPOTGQNTKFTPGVITQDGNHQNPEQQWVLPNYSRGTGHNVLHAPA 420
Db 361 KLGSVQYTTDTNNDPOTGQNTKFTPGVITQDGNHQNPEQQWVLPNYSRGTGHNVLHAPA 420
QY 421 VAPTFGEQLLFRSTMPGCGYPNNMLDCLLPQEWVQHFQCEAAQAQSDVALLRFVNDP 480
Db 421 VAPTFGEQLLFRSTMPGCGYPNNMLDCLLPQEWVQHFQCEAAQAQSDVALLRFVNDP 480
QY 481 TGRVLFECKLHKSQYVTVTAHTGPHDLVIPPNNGYFRFDSWVNFQYTLAPMNGAGRRAL 539
Db 481 TGRVLFECKLHKSQYVTVTAHTGPHDLVIPPNNGYFRFDSWVNFQYTLAPMNGAGRRAL 539
QY 539
Db 539
RESULT 7
Q9WRZ1
ID Q9WRZ1 PRELIMINARY; PRT; 539 AA.
AC Q9WRZ1;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus isolates.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OC NCBI_taxid=150080;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=345-2/96002737/1996/SC;
RX MEDLINE=99246332; PubMed=10228052;
RA Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;
RT "Identification of a distinct common strain of 'Norwalk-like viruses' having a global distribution."
RL J. Infect. Dis. 179:1334-1344 (1999).
DR EMBL; AF080550; AAD40489.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58875 MW; 072C697DD93F3BA7 CRC64;
Query Match 99.3%; Score 2867; DB 12; Length 539;
Best Local Similarity 99.4%; Pred. No. 1.2e-214;
Matches 536; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKMASNDANPSDGSSTANLVPEVNEVMALPEVVGAAIAAPVAGQONVIDPWRNNFVOAP 60
Db 1 MKMASNDANPSDGSSTANLVPEVNEVMALPEVVGAAIAAPVAGQONVIDPWRNNFVOAP 60
QY 61 GGEFTVSPRNAGEILWSAPLGPDLNPLYSLHARMYNGYAGFEQVILAGNAFTAGKII 120
Db 61 GGEFTVSPRNAGEILWSAPLGPDLNPLYSLHARMYNGYAGFEQVILAGNAFTAGKII 120
QY 121 FAAVPPNPTTEGLSPSQVTMPFHIIIDVRQLEPVLIPLDVVRNNFYHYNQSDSTIKLIA 180
Db 121 FAAVPPNPTTEGLSPSQVTMPFHIIIDVRQLEPVLIPLDVVRNNFYHYNQSDSTIKLIA 180
QY 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDDFDIFLVPPTVESRTKPTFTVPIITVEEMSN 240
Db 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDDFDIFLVPPTVESRTKPTFTVPIITVEEMSN 240
QY 241 RFPPIPLEKLYTGPSAFVVPQNGRCTTGGVLLGTTQLSAVNICTRFGDVTTHIAGSHDYT 300
Db 241 RFPPIPLEKLYTGPSAFVVPQNGRCTTGGVLLGTTQLSAVNICTRFGDVTTHIAGSHDYT 300
```



```
QY 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVHTP 360
Db 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVHTP 360
QY 361 KLGSGVQYTTDTNDFQTGNTKFTPVGVIQDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
Db 361 KLGSGVQYTTDTNDFQTGNTKFTPVGVIQDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
QY 421 VAPTFPGEQLLFFRSTMPGCGSGYPNNLDCLLPQEWVQHFCEAAAPQSDVALLRFVNDP 480
Db 421 VAPTFPGEQLLFFRSTMPGCGSGYPNNLDCLLPQEWVQHFCEAAAPQSDVALLRFVNDP 480
QY 481 TGRVLFECKLHSGYVTVHAHTGPHDLVTPPNNGYFRFDSWVNOFYTLAPMNGAGRRRAL 539
Db 481 TGRVLFECKLHSGYVTVHAHTGPHDLVTPPNNGYFRFDSWVNOFYTLAPMNGAGRRRAL 539

RESULT 8
QWRY9
ID QWRY9 PRELIMINARY; PRT; 539 AA.
AC QWRY9;
DT 01-NOV-1999 (TremBLrel. 12, Created)
DT 01-NOV-1999 (TremBLrel. 12, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus isolates.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=150080;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=366/96019554/1996/ID;
RX MEDLINE=99246332; PubMed=10228052;
RA Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;
RT "Identification of a distinct common strain of 'Norwalk-like viruses'
   having a global distribution.";
RL J. Infect. Dis. 179:1334-1344 (1999).
DR EMBL; AF080554; AAD40493.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat_1; -.
SQ SEQUENCE 539 AA; 58842 MW; C7B31E3126A8F4FF CRC64;

Query Match 99.3%; Score 2865; DB 12; Length 539;
Best Local Similarity 99.3%; Pred. No. 1.7e-214;
Matches 535; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
Db 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
QY 61 GGEFTVSRNAPGELTWSAPLGPDLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
Db 61 GGEFTVSRNAPGELTWSAPLGPDLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
QY 121 FAAPVPPNPTGELSPSQVTMFPHIIVDVRLQLEPVLIPLDVRNNFYHYNQSDSTIKLIA 180
Db 121 FAAPVPPNPTGELSPSQVTMFPHIIVDVRLQLEPVLIPLDVRNNFYHYNQSDSTIKLIA 180
QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPVTESRTKPTFTVILTVEMSGNS 240
Db 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPVTESRTKPTFTVILTVEMSGNS 240
QY 241 RPIPLEKLYTGPSSAFVVPQNGRCTTDLGLTQTLQSAVNICFRGDVTHIAGSHDYT 300
Db 241 RPIPLEKLYTGPSSAFVVPQNGRCTTDLGLTQTLQSAVNICFRGDVTHIAGSHDYT 300
QY 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVHTP 360
Db 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVHTP 360
QY 361 KLGSGVQYTTDTNDFQTGNTKFTPVGVIQDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
Db 361 KLGSGVQYTTDTNDFQTGNTKFTPVGVIQDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
```

```
Db 361 KLGSGVQYTTDTNDFQTGNTKFTPVGVIQDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
QY 421 VAPTFPGEQLLFFRSTMPGCGSGYPNNLDCLLPQEWVQHFCEAAAPQSDVALLRFVNDP 480
Db 421 VAPTFPGEQLLFFRSTMPGCGSGYPNNLDCLLPQEWVQHFCEAAAPQSDVALLRFVNDP 480
QY 481 TGRVLFECKLHSGYVTVHAHTGPHDLVTPPNNGYFRFDSWVNOFYTLAPMNGAGRRRAL 539
Db 481 TGRVLFECKLHSGYVTVHAHTGPHDLVTPPNNGYFRFDSWVNOFYTLAPMNGAGRRRAL 539

RESULT 9
QWRY9
ID QWRY9 PRELIMINARY; PRT; 539 AA.
AC QWRY9;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Berlin/495/00/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=173934;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Berlin/495/00/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
   with Norwalk-like viruses in Germany.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBSJ databases.
DR EMBL; AF427123; AAL18879.1; -.
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat_1; -.
SQ SEQUENCE 539 AA; 58869 MW; BC145CA144699DAC CRC64;

Query Match 99.2%; Score 2863; DB 12; Length 539;
Best Local Similarity 99.3%; Pred. No. 2.4e-214;
Matches 535; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
Db 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
QY 61 GGEFTVSRNAPGELTWSAPLGPDLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
Db 61 GGEFTVSRNAPGELTWSAPLGPDLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
QY 121 FAAPVPPNPTGELSPSQVTMFPHIIVDVRLQLEPVLIPLDVRNNFYHYNQSDSTIKLIA 180
Db 121 FAAPVPPNPTGELSPSQVTMFPHIIVDVRLQLEPVLIPLDVRNNFYHYNQSDSTIKLIA 180
QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPVTESRTKPTFTVILTVEMSGNS 240
Db 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPPVTESRTKPTFTVILTVEMSGNS 240
QY 241 RPIPLEKLYTGPSSAFVVPQNGRCTTDLGLTQTLQSAVNICFRGDVTHIAGSHDYT 300
Db 241 RPIPLEKLYTGPSSAFVVPQNGRCTTDLGLTQTLQSAVNICFRGDVTHIAGSHDYT 300
QY 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVHTP 360
Db 301 MNLASQNSNYDPTBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVHTP 360
QY 361 KLGSGVQYTTDTNDFQTGNTKFTPVGVIQDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
Db 361 KLGSGVQYTTDTNDFQTGNTKFTPVGVIQDGNHNEPQQWVLPNYSGRTHNVHLAPA 420
QY 421 VAPTFPGEQLLFFRSTMPGCGSGYPNNLDCLLPQEWVQHFCEAAAPQSDVALLRFVNDP 480
Db 421 VAPTFPGEQLLFFRSTMPGCGSGYPNNLDCLLPQEWVQHFCEAAAPQSDVALLRFVNDP 480
```

QY 481 TGRVLFPECKLHKS... 539  
Db 481 TGRVLFPECKLHKS... 539

RESULT 10

Q916E8 PRELIMINARY; PRT; 539 AA.  
AC Q916E8; (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus NLV/Altenkirchen 140/01/DE.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=173918;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NLV/Altenkirchen 140/01/DE;  
RA Kuenkel U., Hoehne M., Schreier E.;  
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF425765; AAL18858.1; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral Cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 539 AA; 5829 MW; 06A596B3A7D83B34 CRC64;

Query Match 99.2%; Score 2862; DB 12; Length 539;  
Best Local Similarity 99.1%; Pred. No. 2.8e-214;  
Matches 534; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGS... 60  
Db 1 MKMASNDANPSDGS... 60  
QY 61 GGEFTVSPRNAPG... 120  
Db 61 GGEFTVSPRNAPG... 120  
QY 121 FAAPVPPNPTGSL... 180  
Db 121 FAAPVPPNPTGSL... 180  
QY 181 MLYTPLRANNAGD... 240  
Db 181 MLYTPLRANNAGD... 240  
QY 241 RFPVPLEKLYTGP... 300  
Db 241 RFPVPLEKLYTGP... 300  
QY 301 MNLASQNSWVDP... 360  
Db 301 MNLASQNSWVDP... 360  
QY 361 KLGSVQYTTDND... 420  
Db 361 KLGSVQYTTDND... 420  
QY 421 VAPTFPGQLLFF... 480  
Db 421 VAPTFPGQLLFF... 480  
QY 481 TGRVLFPECKLHKS... 539  
Db 481 TGRVLFPECKLHKS... 539

RESULT 11

Q916E7

Q916E7 PRELIMINARY; PRT; 539 AA.  
AC Q916E7; (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus NLV/Dillingen 259/01/DE.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=173919;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NLV/Dillingen 259/01/DE;  
RA Kuenkel U., Hoehne M., Schreier E.;  
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF425766; AAL18859.1; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral Cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 539 AA; 5855 MW; DC1EC814C684E131 CRC64;

Query Match 99.1%; Score 2860; DB 12; Length 539;  
Best Local Similarity 99.1%; Pred. No. 4.1e-214;  
Matches 534; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGS... 60  
Db 1 MKMASNDANPSDGS... 60  
QY 61 GGEFTVSPRNAPG... 120  
Db 61 GGEFTVSPRNAPG... 120  
QY 121 FAAPVPPNPTGSL... 180  
Db 121 FAAPVPPNPTGSL... 180  
QY 181 MLYTPLRANNAGD... 240  
Db 181 MLYTPLRANNAGD... 240  
QY 241 RFPVPLEKLYTGP... 300  
Db 241 RFPVPLEKLYTGP... 300  
QY 301 MNLASQNSWVDP... 360  
Db 301 MNLASQNSWVDP... 360  
QY 361 KLGSVQYTTDND... 420  
Db 361 KLGSVQYTTDND... 420  
QY 421 VAPTFPGQLLFF... 480  
Db 421 VAPTFPGQLLFF... 480  
QY 481 TGRVLFPECKLHKS... 539  
Db 481 TGRVLFPECKLHKS... 539

RESULT 12

Q915D1 PRELIMINARY; PRT; 539 AA.  
AC Q915D1; (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus NLV/Frankfurt (Oder)/170/99/DE.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OX NCBI\_TaxID=173926;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NLV/Frankfurt (Oder)/170/99/DE;  
 RA Kuenkel U., Hoehne M., Schreier E.;  
 RT "Molecular epidemiology of outbreaks of gastroenteritis associated  
 with Norwalk-like viruses in Germany";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF427114; AAL18870.1; -;  
 DR InterPro: IPR004005; Calici\_coat.  
 DR InterPro: IPR008975; Viral\_Cap\_coat.  
 DR Pfam: PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 539 AA; 58815 MW; 09E2BD6C2B984090 CRC64;

Query Match 99.1%; Score 2859; DB 12; Length 539;  
 Best Local Similarity 99.1%; Pred. No. 4,9e-214;  
 Matches 534; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGTANLVPEVNEVMALPVPVGAIAAPVAGQQNVDPWIRNRFVQAP 60  
 DB 1 MKMASNDANPSDGTANLVPEVNEVMALPVPVGAIAAPVAGQQNVDPWIRNRFVQAP 60  
 QY 61 GGEFTVSPRNAPGELWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120  
 DB 61 GGEFTVSPRNAPGELWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120  
 QY 121 FFAVPPNPTEGLSPSQVTMPFPHIIVDVRLQLEPVLIPLDVRRNFFHYNQSNDSITKILIA 180  
 DB 121 FFAVPPNPTEGLSPSQVTMPFPHIIVDVRLQLEPVLIPLDVRRNFFHYNQSNDSITKILIA 180  
 QY 181 MLYTLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFTVPILTVEMSN 240  
 DB 181 MLYTLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFTVPILTVEMSN 240  
 QY 241 RPPPLEKLYTGPSSAFVQVQNGRCTTGDVLLGTQLTSVAVNICFRGDVTHIAGSHDYT 300  
 DB 241 RPPPLEKLYTGPSSAFVQVQNGRCTTGDVLLGTQLTSVAVNICFRGDVTHIAGSHDYT 300  
 QY 301 MNLASQNNNDYPTBEIPAPLGTDPFVGKIQGLTQTTREDGSTRAHKATVSTGSHVHTP 360  
 DB 301 MNLASQNNNDYPTBEIPAPLGTDPFVGKIQGLTQTTREDGSTRAHKATVSTGSHVHTP 360  
 QY 361 KLGVSQYVTTDNDTQNTKFTPGVVIQGNHQNHPQVWLPNYSGRTGHNVLHAPA 420  
 DB 361 KLGVSQYVTTDNDTQNTKFTPGVVIQGNHQNHPQVWLPNYSGRTGHNVLHAPA 420  
 QY 421 VAPTPGQQLLFFRSTMPGCSGYPNMNLDCLLPQEWVQHFQEAAPASQDVALLRFVNP 480  
 DB 421 VAPTPGQQLLFFRSTMPGCSGYPNMNLDCLLPQEWVQHFQEAAPASQDVALLRFVNP 480  
 QY 481 TGRVLFECKLHKSQVTVTAHTGPHDLVIPPNGYFRFDSWVNOFYTLAPMNGAGRRAL 539  
 DB 481 TGRVLFECKLHKSQVTVTAHTGPHDLVIPPNGYFRFDSWVNOFYTLAPMNGAGRRAL 539

RESULT 13  
 Q91323 ID Q91323 PRELIMINARY; PRT; 539 AA.  
 AC Q91323  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Capsid protein.  
 OS Human calicivirus NLV/VA98387/1998.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OX Norovirus.  
 OX NCBI\_TaxID=165511;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NLV/VA98387/1998;  
 RA Jiang X., Zhong W.M., Farkas T., Huang P.W., Wilton N., Barrett E.,

RA Fulton D., Morrow R., Matson D.O.;  
 RT "Baculovirus expression and antigenic characterization of the capsid  
 proteins of three Norwalk-like viruses";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AV038600; AAK84679.1; -;  
 DR InterPro: IPR004005; Calici\_coat.  
 DR InterPro: IPR008975; Viral\_Cap\_coat.  
 DR Pfam: PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 539 AA; 58887 MW; E9A784B3604C83AC CRC64;

Query Match 99.0%; Score 2857; DB 12; Length 539;  
 Best Local Similarity 98.9%; Pred. No. 7e-214;  
 Matches 533; Conservative 4; Mismatches 2; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGTANLVPEVNEVMALPVPVGAIAAPVAGQQNVDPWIRNRFVQAP 60  
 DB 1 MKMASNDANPSDGTANLVPEVNEVMALPVPVGAIAAPVAGQQNVDPWIRNRFVQAP 60  
 QY 61 GGEFTVSPRNAPGELWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120  
 DB 61 GGEFTVSPRNAPGELWSAPLGPDLNPLYSLHARMYNGYAGFEVQVILAGNAFTAGKII 120  
 QY 121 FFAVPPNPTEGLSPSQVTMPFPHIIVDVRLQLEPVLIPLDVRRNFFHYNQSNDSITKILIA 180  
 DB 121 FFAVPPNPTEGLSPSQVTMPFPHIIVDVRLQLEPVLIPLDVRRNFFHYNQSNDSITKILIA 180  
 QY 181 MLYTLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFTVPILTVEMSN 240  
 DB 181 MLYTLRANAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFTVPILTVEMSN 240  
 QY 241 RPPPLEKLYTGPSSAFVQVQNGRCTTGDVLLGTQLTSVAVNICFRGDVTHIAGSHDYT 300  
 DB 241 RPPPLEKLYTGPSSAFVQVQNGRCTTGDVLLGTQLTSVAVNICFRGDVTHIAGSHDYT 300  
 QY 301 MNLASQNNNDYPTBEIPAPLGTDPFVGKIQGLTQTTREDGSTRAHKATVSTGSHVHTP 360  
 DB 301 MNLASQNNNDYPTBEIPAPLGTDPFVGKIQGLTQTTREDGSTRAHKATVSTGSHVHTP 360  
 QY 361 KLGVSQYVTTDNDTQNTKFTPGVVIQGNHQNHPQVWLPNYSGRTGHNVLHAPA 420  
 DB 361 KLGVSQYVTTDNDTQNTKFTPGVVIQGNHQNHPQVWLPNYSGRTGHNVLHAPA 420  
 QY 421 VAPTPGQQLLFFRSTMPGCSGYPNMNLDCLLPQEWVQHFQEAAPASQDVALLRFVNP 480  
 DB 421 VAPTPGQQLLFFRSTMPGCSGYPNMNLDCLLPQEWVQHFQEAAPASQDVALLRFVNP 480  
 QY 481 TGRVLFECKLHKSQVTVTAHTGPHDLVIPPNGYFRFDSWVNOFYTLAPMNGAGRRAL 539  
 DB 481 TGRVLFECKLHKSQVTVTAHTGPHDLVIPPNGYFRFDSWVNOFYTLAPMNGAGRRAL 539

RESULT 14  
 Q917W2 ID Q917W2 PRELIMINARY; PRT; 539 AA.  
 AC Q917W2  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)  
 DE Capsid protein.  
 GN ORF2.  
 OS Norwalk-like virus NLV/Burwash Landing/331/1995/US.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OX Norovirus.  
 OX NCBI\_TaxID=171857;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/Burwash Landing/331/1995/US;  
 RX MEDLINE=97193806; PubMed=9041391;  
 RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
 a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
 of small round-structured viruses (Norwalk-like viruses).";  
 RL J. Clin. Microbiol. 35:570-577(1997).

```
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Burwash Landing/331/1995/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RA "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Burwash Landing/331/1995/US;
RX MEDLINE=9246332; PubMed=10228052;
RA Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;
RT "Identification of a distinct common strain of 'Norwalk-like viruses'
RT having a global distribution.";
RL J. Infect. Dis. 179:1334-1344 (1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Burwash Landing/331/1995/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414425; AAL13028.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58791 MW; 57A9576B0076B5BC CRC64;

Query Match 99.0%; Score 2856; DB 12; Length 539;
Best Local Similarity 98.9%; Pred. No. 8.3e-214;
Matches 533; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
Db 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
QY 61 GGEFTVSPRNAPGEILLWSAPLGLPDLNPLYSLHARMYNGVAGGEVQVILAGNAFTAGKII 120
Db 61 GGEFTVSPRNAPGEILLWSAPLGLPDLNPLYSLHARMYNGVAGGEVQVILAGNAFTAGKII 120
QY 121 FFAVPPNFPTEGLSPSQVTMFPHIIVDVROLEPVLPLPDVRNNFYHYNQSDNDSTIKLIA 180
Db 121 FFAVPPNFPTEGLSPSQVTMFPHIIVDVROLEPVLPLPDVRNNFYHYNQSDNDSTIKLIA 180
QY 181 MLYTPLRANNAGDDVFTVSCRVLTRPSPDFDFLVPPTVESRTKPTFVILTVEMSN 240
Db 181 MLYTPLRANNAGDDVFTVSCRVLTRPSPDFDFLVPPTVESRTKPTFVILTVEMSN 240
QY 241 RFPIPLEKLYTGPSAFAVQONGRCCTDGVLLGTTQLSANNICTRPGDVTHIAGSHDYT 300
Db 241 RFPIPLEKLYTGPSAFAVQONGRCCTDGVLLGTTQLSANNICTRPGDVTHIAGSHDYT 300
QY 301 MNLASQNSVNDPTEEPAPLGTDFVGGKIQCMLTOTTRDGGSTRAHKATVSTGSHFTP 360
Db 301 MNLASQNSVNDPTEEPAPLGTDFVGGKIQCMLTOTTRDGGSTRAHKATVSTGSHFTP 360
QY 361 KLGSVQVTTDNTNDPQTGNTKFTFVGVIGQNNHNEPQGMVLNPNYSGRTGHNVLAPA 420
Db 361 KLGSVQVTTDNTNDPQTGNTKFTFVGVIGQNNHNEPQGMVLNPNYSGRTGHNVLAPA 420
QY 421 VAPTFPGEQLLFFRSTWEGCSGYNNMNDCLLPQEWQHFCEAQAQSDVALLRFVNP 480
Db 421 VAPTFPGEQLLFFRSTWEGCSGYNNMNDCLLPQEWQHFCEAQAQSDVALLRFVNP 480
QY 481 TGRVLFECKLHKSQVTVVAHTGPHDLVTPNGYFRFDSWNQFYTLAPMGNGAGRRAL 539
Db 481 TGRVLFECKLHKSQVTVVAHTGPHDLVTPNGYFRFDSWNQFYTLAPMGNGAGRRAL 539

RESULT 15
Q917W5
```

```
ID Q917W5 PRELIMINARY; PRT; 539 AA.
AC Q917W5;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Miami Beach/326/1995/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_taxid=171856;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Miami Beach/326/1995/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577 (1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Miami Beach/326/1995/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383 (1997).
RN [3]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Miami Beach/326/1995/US;
RX MEDLINE=9246332; PubMed=10228052;
RA Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;
RT "Identification of a distinct common strain of 'Norwalk-like viruses'
RT having a global distribution.";
RL J. Infect. Dis. 179:1334-1344 (1999).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Miami Beach/326/1995/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348 (2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Miami Beach/326/1995/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414424; AAL13025.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58759 MW; D6A89AF9269DCFD CRC64;

Query Match 99.0%; Score 2856; DB 12; Length 539;
Best Local Similarity 99.1%; Pred. No. 8.3e-214;
Matches 534; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
Db 1 MKMASNDANPSDGSSTANLVPEVNEVMALEPVVGAIAAPVAGQQNVDPWIRNNFVQAP 60
QY 61 GGEFTVSPRNAPGEILLWSAPLGLPDLNPLYSLHARMYNGVAGGEVQVILAGNAFTAGKII 120
Db 61 GGEFTVSPRNAPGEILLWSAPLGLPDLNPLYSLHARMYNGVAGGEVQVILAGNAFTAGKII 120
QY 121 FFAVPPNFPTEGLSPSQVTMFPHIIVDVROLEPVLPLPDVRNNFYHYNQSDNDSTIKLIA 180
Db 121 FFAVPPNFPTEGLSPSQVTMFPHIIVDVROLEPVLPLPDVRNNFYHYNQSDNDSTIKLIA 180
```

Qy	181	MLYTPLRANNAGDDVFTVSCRVLTRPSPDFIFLVPPTVESRTKPF	240
Db	181	MLYTPLRANNAGDDVFTVSCRVLTRPSPDFIFLVPPTVESRTKPF	240
Qy	241	REPIPLEKLYTGPSSAFVVPQNGRCTTDGVLGGTTQLSAVNICTFRG	300
Db	241	REPIPLEKLYTGPSSAFVVPQNGRCTTDGVLGGTTQLSAVNICTFRG	300
Qy	301	MNLASQWNSNYDPTTEEIPAPLGTDPFVGKIQQMLTQTTREDGSTR	360
Db	301	MNLASQWNSNYDPTTEEIPAPLGTDPFVGKIQQMLTQTTREDGSTR	360
Qy	361	KLGSVQYTTDTNDFOTGQNTKFTPVGVIQDGNHONEPQWVLPNYS	420
Db	361	KLGSVQYTTDTNDFOTGQNTKFTPVGVIQDGNHONEPQWVLPNYS	420
Qy	421	VAPTFPEGOLLFRSTMPGCSGYPNMNLDCLLPQEWVQHFCOEAAP	480
Db	421	VAPTFPEGOLLFRSTMPGCSGYPNMNLDCLLPQEWVQHFCOEAAP	480
Qy	481	TGRVLFECKLHKSGYVYVAHTGPHDLVIPPNGYFRFDSWVNOFYTL	539
Db	481	TGRVLFECKLHKSGYVYVAHTGPHDLVIPPNGYFRFDSWVNOFYTL	539

Search completed: June 1, 2004, 13:53:16  
Job time : 34.3814 secs

**THIS PAGE BLANK (USPTO)**



QY	181	MLYTPLRANNAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTKPTFTVPLTVEEMSNS	240
Db	181		
QY	241	RPPIPLEKLYTGPSSAFVQPNQGRCTTGDVLLGTTLQTSANVICTFRGDVTHIAGSHDYT	300
Db	241		
QY	301	MLNASQNSNYDPTTEIPAPLCTGPDVGVKIQGLTQTTRDGSTRAHKATVSTGSHVFTP	360
Db	301		
QY	361	KLGSVOYTTDNNDFOTGQNTKFTPVGVIGDGNHQNHPQQWVLPNYSGRTGHNHVLAPA	420
Db	361		
QY	421	VAPTFPGEQLLFFRSTMPGCGYPNNLDCLLPQEWVLFYQEAAPQSDVALLRFVNP	480
Db	421		
QY	481	TGRVLECKLHKSYYTVTAHTGPHDLVIPPNGYFRFDSWVNOFYTLAPMNGAGRRRAL	539
Db	481		
RESULT 2			
ID	COAT_SOUV3	STANDARD;	PRT; 546 AA.
AC	Q04542;		
DT	01-OCT-1994 (Rel. 30, Created)		
DT	01-OCT-1994 (Rel. 30, Last sequence update)		
DT	01-OCT-1994 (Rel. 30, Last annotation update)		
DE	Coat protein (Capsid protein).		
OS	Southampton virus (serotype 3).		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_TaxID=37129;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
EX	MEDLINE=93142023; PubMed=8380940;		
RA	Lambden P.R.; Caul E.O.; Ashley C.R., Clarke I.N.;		
RT	"Sequence and genome organization of a human small round-structured		
RT	(Norwalk-like) virus.";		
RL	Science 259:516-519(1993).		
CC	-1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.		
CC	-----		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	-----		
DR	EMBL; L07418; AAA92984.1; --		
DR	PIR; B37491; B37491.		
DR	InterPro; IPR004005; Calici_coat.		
DR	InterPro; IPR008975; Viral_Cap_coat.		
DR	Pfam; PF00915; Calici_coat; 1.		
KW	Coat protein; Glycoprotein.		
FT	CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).		
SQ	SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;		
Query Match			
Best Local Similarity 44.6%; Pred No. 1.6e-74;			
Matches 249; Conservative 80; Mismatches 193; Indels 36; Gaps 17;			
QY	1	MKMASNDANPFS-DGST--ANLVPEVNN-EVMALEPVVGAAIAAPVAGQNVLDPIRNVF	56
		: : : : :       : : : : :       : : : : :	



```
CC EMBL; M67473; AAA47285.1; --
DR PIR; A41039; RRMWRH.
DR PDB; 1KHV; 16-JAN-02.
DR PDB; 1KHV; 16-JAN-02.
DR MEROPS; C24.001; --
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR004005; Calci_coat.
DR InterPro; IPR004004; Calci_pol_hel.
DR InterPro; IPR000317; Peptidase_C24.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR Pfam; PF03510; Peptidase_C24; 1.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PRO0916; 2CENDOPTASE.
DR PRINTS; PRO0918; CALICIVIRUSNS.
DR SMART; SM00382; AAA; 1.
DR PolyProtein; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.
FT CHAIN 7508 7707 HELICASE (P2C-LIKE).
FT CHAIN 1109 1251 PROTEASE P3C.
FT CHAIN 71497 71625 RNA-DIRECTED RNA POLYMERASE.
FT NP BIND 522 529 ATP (POTENTIAL).
FT ACT_SITE 1212 1212 PROTEASE (POTENTIAL).
FT ACT_SITE 1227 1227 PROTEASE (POTENTIAL).
SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 11.3%; Score 326; DB 1; Length 2344;
Best Local Similarity 25.3%; Pred. No. 1e-15;
Matches 137; Conservative 66; Mismatches 209; Indels 130; Gaps 25;

QY 4 ASNDANPSDGTANILVPEVNNVMALEPVVGAIAAPVAGQONVID---PWIRNNFVQAP 60
Db 1785 ASVPGTTTDMGPGVV--ATTSVTAENSSASIAATAGIGPPQVDDQETW-RTNFYY-- 1839

QY 61 GEEFTVSPNARGEILMSAPGLDNLNVLHARMVYAGGFEVQVILAGNAFTAGKII 120
Db 1840 NDVFTWSVADAGSILYTVQHSPPNNPFTAVLSQYAGWAGMGQFRFVAGSGVGGRLV 1899

QY 121 FAAPVPNPFTEGLSPS-QVTMPHIIIVDVRQLEPVLIPLDVRRNFYHYNQSDSTIKLI 179
Db 1900 RAVIIPGIE---IGGLEVRQFPHVIDARSLEPVTIIMPDLRPNMYHTGDPGLVPTLV 1956

QY 180 AMLYPLRANNAGDDVFTVSCRVLTRPSPDFIFLVP--TVESRTKP--FTVPILTV 234
Db 1957 LSVYNNL-INPFGGSTSAIQVTVETRPSEDPEFVMIRAPSSKTVDISPAGLTTPVLTL- 2014

QY 235 EEMSNRPPIPLEKLYTGSSAFVQPNQ---RCTTDGLVLTGTTQ-LSAVNICTRGDV 290
Db 2015 -----GVGNDRWNG--QIVGLQPVPGFSTCNHRHNLNGSTYGWSPRF-----GDI 2060

QY 291 THIAGSHDYTMNLAS---QNW-----SNVD-PTEIIPAPLGTDP----- 325
Db 2061 DHRGASYSGSNATNVLQFVYANAGSAIDNPISQV-APDGPDMFSVPFNGPGGIPAAGW 2119

QY 326 -----FVGKIQMLTQTTREDGSTRAHKA--TVSTGS--- 355
Db 2120 VGFGLIWNNSGAPNVTTVQAVELGFATGAPGNLQPTINTSGAQTVAKSIYAVVTGTAQN 2179

QY 356 -----VHFTPKLSVQVTTTNDNFQT-----GQNTKTFPVGVIQ----- 390
Db 2180 PAGLFVMSAGIISTNASAITTPQDRIVTTTGTPTAAAPVAKNTPIMFASVYRRTGDVN 2239

QY 391 ----DGNHQNPEQQWLPNYSGRTHNVHAPAVAPTFGQLLFFRSTMPGCSGYPNM 446
Db 2240 ATAGSANGTYGTGSOPLPVTIIGLSNNYSSA-----LMPGQ-----FFWQLTTFASGFMEI 2291
```

```
QY 447 NL 448
Db 2292 GL 2293

RESULT 4
COAT_FCVF9 STANDARD; PRT; 671 AA.
AC P27406;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Coat protein (Capsid protein).
GN CYP76.
OS Feline calicivirus (strain F9) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11981;
RN [1]_TaxID=11981;
RP SEQUENCE FROM N.A.
RX MEDLINE=92410623; PubMed=1529544;
RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,
RA Turner P.C.;
RT "The complete nucleotide sequence of a feline calicivirus.";
RL Virology 190:443-448(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339673; PubMed=1633955;
RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;
RT "Cloning and sequence determination of the feline calicivirus strain
RT F9.";
RL Biochem. Soc. Trans. 20:265-265(1992).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL; M86379; AAA79327.1; --
DR EMBL; Z11536; CAA77636.1; --
DR PIR; B43382; VCMWPF9.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 671 AA; 73441 MW; 33BE86D8370D5ES CRC64;

Query Match 10.4%; Score 300.5; DB 1; Length 671;
Best Local Similarity 25.3%; Pred. No. 1.4e-14;
Matches 141; Conservative 80; Mismatches 238; Indels 99; Gaps 28;

QY 12 DGSITANLVPEVNNV--MALEPVVGAIAAPVAGQONVIDPWIRNNFVQAPGGEFTV--- 66
Db 127 DGSIT--APEQGTVMVGVIAPESQAQMSTAADATKGSVDSEW-----EAPFSPTSVNW 178

QY 67 SPRNAPCEILASAPLDNLNVLHARMVYAGGFEVQVILAGNAFTAGKIIFAAVPP 126
Db 179 STSETQKILPKQSLGSLNPLNLEHLAKLYVWAGSIEVRFSISGSGVFGKLAIVVPP 238

QY 127 NFPTEGSPSQVT---MPPHIIIVDVRQLEPVLIPLDVRRNFYHYNQSDSTIKLIAMLY 183
Db 239 -----GVDVPQSTMLQYPHVLFDAQVEPVIFCLPDLRLSTLYHLMSDIT-SLIVMY 292

QY 184 TPLRANNAGDDVFTVSC--RVLTRSPDPDFIFLVPPTVESRTRKPTFTVILTVEMSNSR 241
```



```

DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 339 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 668 AA; 73589 MW; 85BDBC85804E503 CRC64;

Query Match 10.2%; Score 294; DB 1; Length 668;
Best Local Similarity 24.6%; Pred. No. 4.2e-14;
Matches 120; Conservative 61; Mismatches 171; Indels 136; Gaps 21;

QY 12 DGSTANLVPEVNEV--MALEPVVGAATAAPVAGQONVIDPWIRNPNVQAPGGEFTV--- 66
DB 127 DGSIT--TPEQGTAVGVIAEPAQMSAQMSTADWATKSKVDSEW-----EAPFSFHTSNVW 178

QY 67 SPRNAPGEILMSAPLGPDLPNLYSLHARMYNGYAGGFEVQVILAGNAFTAGKIIFAAVPP 126
DB 179 STSETQKILFKQSLGPLLPNLYSLHARMYNGYAGGFEVQVILAGNAFTAGKIIFAAVPP 238

QY 127 NPTTEGLSPSQVT---MPFHIIVDVRQLEPVLPLPDVRRNPNFYHYNQNSDSTIKLIAMLY 183
DB 239 -----GVDPVQSTSLMQLQPHVLFDAQVEPFITPDLRSLRSTLYHVMSTDTT-SLVIMVY 292

QY 184 TPLRANNAGDDVFTVSC--RVLTSPSPDFEFLVPP-----TVES----- 222
DB 293 NDL--INPYANDNSGCIIVTETKPGDPKFHLLKPPGSLVTHGSIPLSDLPKSSSLWIG 351

QY 223 -----RTKPFVTPILTVEEMSNRFPPIPLE 247
DB 352 NRYWTDITDFVIRPFVQANRHFDFNQETAGSTFRPRITI--TISEKNSKLGIGVA 408

QY 248 KLYTGP-----SSAFVQPNQGRCTTGDVLLGTQ---LSAV--NICTPRGDTV 291
DB 409 TDYIIFGIPDGWPDFTTIADKLIPAGDYSITTTGEGNDIKTAQAYDTAAVYVNTTFRG--M 466

QY 292 HTAGSHDYTMNLASQNSNYDTEIPAPLGTGPDFFVGKIQGLMLOTTR-EDGSTRAHKAT 350
DB 467 YICGS-----LQAWGD-----KKISNTAFITTAIRDGNEIKPSNT 502

QY 351 VSTGSHVHTPKLGSQYVYTTDNTNDFQTGONT-----KFTPVGVIOGNNHQPQWVLPN 406
DB 503 IDMTKL-----AVYQDTHVEQEVQTSDDTLALLGYTGIGEEAIGSNRDRVVRISVLPE 555

QY 407 YSGRTGHN 414
DB 556 AGARGGNH 563

RESULT 7
COAT_FCV66 STANDARD; PRT; 668 AA.
AC P27404;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
OS Coat protein (Capsid protein).
OS Feline calicivirus (strain CFI/68 FIV) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11979;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91374597; PubMed=1716692;
RA Neill J.D., Reardon I.M., Heinrichson R.L.;
RT "Nucleotide sequence and expression of the capsid protein gene of
RL feline calicivirus.";
RN J. Virol. 65:5440-5447(1991).
[2]

SEQUENCE FROM N.A.
Neill J.D.;
Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
-!- PTM: The N-terminus is blocked.
-!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL; M32819; AAA42925.1; -.
CC EMBL; U13992; AAC13993.1; -.
DR PIR; A40507; VCMWFF.
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 439 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 668 AA; 73550 MW; 9E52312108D4D41 CRC64;

Query Match 10.2%; Score 293; DB 1; Length 668;
Best Local Similarity 25.3%; Pred. No. 5e-14;
Matches 117; Conservative 66; Mismatches 195; Indels 84; Gaps 18;

QY 12 DGSTANLVPEVNEV--MALEPVVGAATAAPVAGQONVIDPWIRNPNVQAPGGEFTV--- 66
DB 127 DGSIT--TPEQGTAVGVIAEPAQMSAQMSTADWATKSKVDSEW-----EAPFSFHTSNVW 178

QY 67 SPRNAPGEILMSAPLGPDLPNLYSLHARMYNGYAGGFEVQVILAGNAFTAGKIIFAAVPP 126
DB 179 STSETQKILFKQSLGPLLPNLYSLHARMYNGYAGGFEVQVILAGNAFTAGKIIFAAVPP 238

QY 127 NPTTEGLSPSQVT---MPFHIIVDVRQLEPVLPLPDVRRNPNFYHYNQNSDSTIKLIAMLY 183
DB 239 -----GIDPVQSTSLMQLQPHVLFDAQVEPFITPDLRSLRSTLYHVMSTDTT-SLVIMVY 292

QY 184 TPLRANNAGDDVFTVSC--RVLTSPSPDFEFLVPP-----TVES----- 241
DB 293 NDL--INPYANDNSGCIIVTETKPGDPKFHLLKPP--GSMITHGSIPLSDLPKSSS--- 347

QY 242 FPIPLEKLYTGPFSSAFVQVP---QNGR-----CTTDGVLGTTQLSAVNICTFRGDVTHI 293
DB 348 --LWIGNRFWSDITDFVIRPFVQANRHFDFNQETAGSTFRPRITITISVKESAKLGI 405

QY 294 AGSHDYTMNLASQNSNYDTEEP---APLGTDPDFVGKIQGLMLOTTR-EDGST- 345
DB 406 GVATDYIVPGIPDGN---PDTTIPGELVPVGDYAITGTNTDITTAQYDAATEIRNNT 461

QY 346 -----AHKATVSTGSHVHTPKLGS-----VQYTTDNTNDFQ 376
DB 462 NFRGMYICGSLQRAWGDKKISNTAFITGTVDGAKLIPSNITDQTKIAVQDTHANKHVQ 521

QY 377 TQONT----KFTPVGVIOGNNHQPQWVLPNYSGRTHN 414
DB 522 TSDDTLALLGYTGIGEEAIGADRVRVIRISVLPERGARGGNH 563

RESULT 8
COAT_SMSV4
ID COAT_SMSV4 STANDARD; PRT; 703 AA.
AC P36285;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)

```

DE Coat protein (capsid protein).  
OS San Miguel sea lion virus (serotype 4) (SMSV 4).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Vesivirus.  
OX NCBI\_TaxID=36407;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92410750; PubMed=1529644;  
RA Neill J.D.;  
RT "Nucleotide sequence of the capsid protein gene of two serotypes of  
RT San Miguel sea lion virus: identification of conserved and non-  
RT conserved amino acid sequences among calicivirus capsid proteins.";  
RL Virus Res. 24:211-222(1992).  
CC 1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC -----  
CC EMBL; M87482; AAA16220.1; --  
CC PIR; C48562; C48562.  
CC InterPro; IPR004005; Calici\_coat.  
CC Pfam; PF00915; Calici\_coat; 1.  
CC Coat protein; glycoprotein.  
KW CARBOHYD 89 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 208 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 329 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 463 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 482 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 703 AA; 77721 MW; C5DAD823B261073 CRC64;  
  
Query Match 10.8%; Score 288.5; DB 1; Length 703;  
Best Local Similarity 27.3%; Pred. No. 1.1e-13;  
Matches 101; Conservative 56; Mismatches 164; Indels 49; Gaps 16;  
  
QY 11 SDG-STANLVPEVNNVMALEPV---VGAIAAPVAGQNVVIDPWRNNFVQAPGGEFTV 66  
DB 153 SDGSGSEIVTEEGQTVVQQPAPAPATLATATAGTSGVQEQEWMTFFSYHTSNKSTV 212  
  
QY 67 SPRNAPGEILWSAPGLDPLNPNYLHARMNGVAGGEVQVILAGNAFTAGKILPAVPP 126  
DB 213 ---ESQKILYSQALNPSINPLDHIARLYSTWSGGIDVRFVTSGSGVFGKLAALLVPP 269  
  
QY 127 NF-PTEGLSPSOVTMFPHIIVDQLPVLPLPDLVRNNFVHYNQSNDSITIKLIAMLYTP 185  
DB 270 GVEPIESVSMIQ---YPHVLFDARQTEPVFTIPDIRKTLFHSMDTDTT-KLVINPY-- 323  
  
QY 186 LRANNAGDDVTVSCRVLTRPSDFDFILVPP--TVESRTKPF-TVPILTVEMSMNSRF 242  
DB 324 ---ENGVENKTTCSITVETRESADFTFALLKPGSLIKHGSIPSDILPRNSAHWMGNR-- 378  
  
QY 243 PIPLEKLYTGSPSSAFVQPOQ-----NGRCTTGVLGTLTQLSAVNI-CTFRGCVTHIAGSH 297  
DB 379 -----WWSITSGFSVQPRVFSQNRHDFDSTTTGMSSTPYVPIETKIQKV-----GSN 427  
  
QY 298 DYTWNLASQNSNYDPTTEIPALGTDFVKGIOGMLTQTTREDGSTRAHKATYSTGVSFH 357  
DB 428 N-----KMFHVITDOKALVP-GIPD--GWPDITPIDETKATNGNSYGESYRAGSTT 476  
  
QY 358 FTPKLGVSQY 367  
DB 477 IKPNENSTHF 486  
  
RESULT 9  
POLN MANCV STANDARD; PRT; 2208 AA.  
ID POLN MANCV  
AC Q69014;

DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polyprotein [Contains: RNA-directed RNA polymerase  
DE (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like  
DE protein); Coat protein].  
OS Manchester virus (Human enteric calicivirus).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Sapovirus.  
OX NCBI\_TaxID=82659;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95390791; PubMed=7661689;  
RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;  
RT "Human enteric caliciviruses have a unique genome structure and are  
RT distinct from the Norwalk-like viruses.";  
RL Virology. 140:1345-1356(1995).  
CC 1- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).  
CC 1- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR  
CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY  
CC SIMILARITY).  
CC 1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA} (N).  
CC 1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC 1- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.  
CC 1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC -----  
CC EMBL; X86560; CAA60262.1; --  
CC MEROPS; C24.UPW; --  
CC InterPro; IPR004005; Calici\_coat.  
CC InterPro; IPR004004; Calici\_pol\_hel.  
CC InterPro; IPR000317; Peptidase\_C24.  
CC InterPro; IPR000605; RNA\_helicase.  
CC InterPro; IPR007095; RNA\_pol\_DS\_PS.  
CC InterPro; IPR001205; RNA\_pol\_P3D.  
CC InterPro; IPR007094; RNA\_pol\_PSVir.  
CC InterPro; IPR008975; Viral\_cap\_coat.  
CC Pfam; PF00915; Calici\_coat; 1.  
CC Pfam; PF03510; Peptidase\_C24; 1.  
CC Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
CC Pfam; PF00910; RNA\_helicase; 1.  
CC PRINTS; PR00916; 2CENDOPTASE.  
CC PRINTS; PR00918; CALICIVIRUSN.  
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;  
KW Thiol protease; Helicase; ATP-binding; Coat protein.  
FT CHAIN ? ? HELICASE (P2C-LIKE).  
FT CHAIN ? ? PROTEASE (P3C).  
FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.  
FT CHAIN ? ? COAT PROTEIN.  
FT NP\_BIND 408 415 ATP (POTENTIAL).  
FT ACT\_SITE 1097 1097 PROTEASE (BY SIMILARITY).  
FT ACT\_SITE 1112 1112 PROTEASE (BY SIMILARITY).  
SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;  
  
Query Match 9.9%; Score 285; DB 1; Length 2208;  
Best Local Similarity 23.0%; Pred. No. 1e-12;  
Matches 148; Conservative 78; Mismatches 224; Indels 194; Gaps 27;  
  
QY 2 KMASNDANPSDGSSTANLVPEVNNVMALEP-----VV-----GAA-----IAA 39  
DB 1647 EMEGNGSNPE-----PKQNNPMVDPFGTTGTTTSHVVVNPANPEQNGAAQRLAV 1698  
  
QY 40 PVAGQQNVIDPWIRNNFVQAPGGEFTVSPRNAPGEILWSAPGLDPLNPNYLHARMYNCY 99  
DB 1699 ATGAIQSNVPEAIRNCF--AVFRTFAWNRDMPTGTFLGSIHLPNINPYTSLHSGWAGW 1756

QY 100 AGFEVQVVLNAGNATAGKIIFAAVPPNPPTREGLSPSQVT---MPPHIIIVDVRQLEPVLI 156  
 DB 1757 GGSFEVRLSISGSGVAGRIIASVIPP-----GVDPSSIRDQVLPFAFVDARITEPVSF 1811  
 QY 157 PLPDVNNHYHYNQSNDSITIKIAMLYTLPLRANNAGDVFTVSCRVLTSPSPDFDFELV 216  
 DB 1812 MIPDVRADVHRMDGAEPTCSLGFVYQYLLPFPFTTAVSTCWVSVEKPGGDFDFCLLR 1871  
 QY 217 PPTVE-----S 222  
 DB 1872 PPGQMEGVSPGELLPRRLGYSGRNVUGLVVGMILVAEHKQVNRHFNFSNVSFTGWSFA 1931  
 QY 223 RTKPTVPLITVEEMNSR---FPIPLEK---LYTGPSSAFVQPNQGRCTTDGVLGTT 276  
 DB 1932 PYNMAAEIVTNOAHSTSHAWLSICAQNKGLFPFGIPNH---PDSCASTVVGAM--DT 1986  
 QY 277 QLS---AVNIC---TFR---GDVTHIAGSHDVTMLASQNWNSDYPTBEIPAPLPTPDV 327  
 DB 1987 SUGRPSTGCGPAISFQNGDV-----YENDTPSVMFATYDP----- 2024  
 QY 328 GKIQGLTQTTREDGSTRAHKATVSTGSHVFTPKLGSVQYTTDTWNDFQTG--QNTKEFP 385  
 DB 2025 -----LTSGT-----GVALTNSIN-----PASLALVRISNNDPDTSGFANDKNV 2064  
 QY 386 VGVIOQGNHNEPQMWLP-----NYS-GRTHGN-VHLAPAVAPTFPGQLLFFRSTWPG 439  
 DB 2065 VQMSWEMVTGTNIQIRGQVTPMSGTNYTFTSTGANTLVLMQERMLSYDGHQALYSQLER 2124  
 QY 440 CSGYPNMNLCLLPQEWQHFCQEAAPQSDVA-----LLRFVNDTGRVLPCKLHKS 494  
 DB 2125 TAEY-----FQNDIVNIPENSMAVFNVTNSASFQIGIRPDG 2161  
 QY 495 YVTVAHTGPHDLVIPNGYFRFDSVMNQFYTLA-PMGN-GAGRR 536  
 DB 2162 YMTGSGIVNVPLPETFQVVGILPLSAALSGSGNWRARR 2205

## RESULT 10

POLG\_CX16T STANDARD; PRT; 2193 AA.  
 AC Q9QF31.  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].  
 OS Coxsackievirus A16 (strain Taiwan/5079/98).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.  
 OC NCBI\_TaxID=231417;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21427116; PubMed=11536241;  
 RA Yan J.-J., Su I.-J., Chen P.-F., Liu C.-C., Yu C.-K., Wang J.-R.;  
 RT "Complete genome analysis of enterovirus 71 isolated from an outbreak in Taiwan and rapid identification of enterovirus 71 and coxsackievirus A16 by RT-PCR."  
 RL J. Med. Virol. 65:331-339(2001).  
 CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular membranes and is associated with viral RNA synthesis.  
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).  
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2,

CC VP3, and VP4.  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC  
 CC EMBL; AF177911; AAD55085.1; --  
 CC HSSP; P03300; IPOV.  
 CC InterPro; IPR003593; AAA ATPase.  
 CC InterPro; IPR004004; Calici pol hel.  
 CC InterPro; IPR009003; Cys Ser trypsin.  
 CC InterPro; IPR000199; Pept\_3C\_picorn.  
 CC InterPro; IPR000081; Peptidase\_C3.  
 CC InterPro; IPR003138; Pico\_P1A.  
 CC InterPro; IPR002527; Pico\_P2B.  
 CC InterPro; IPR001676; Rhv.  
 CC InterPro; IPR006005; RNA helicase.  
 CC InterPro; IPR007095; RNA pol DS PS.  
 CC InterPro; IPR001205; RNA pol\_P3D.  
 CC InterPro; IPR007094; RNA pol\_PSVir.  
 CC InterPro; IPR008975; Viral cap coat.  
 CC Pfam; PF00548; Cys-protease-3C; 1.  
 CC Pfam; PF02226; Pico\_P1A; 1.  
 CC Pfam; PF00947; Pico\_P2A; 1.  
 CC Pfam; PF01552; Pico\_P2B; 1.  
 CC Pfam; PF00073; rhv; 3.  
 CC Pfam; PF00680; RNA dep RNA pol; 1.  
 CC Pfam; PF00910; RNA helicase; 1.  
 CC PRINTS; PR00918; CALICIVIRUSNS.  
 CC ProDom; PD001125; Cys\_protease\_3C; 1.  
 CC ProDom; PD001306; Pico\_P2A; 1.  
 CC ProDom; PD001274; Pico\_P2B; 1.  
 CC SMART; SM00382; AAA; 1.  
 CC Polyprotein; Coat protein; Core protein; Thiol protease; Myristate;  
 KW RNA-directed RNA polymerase; Hydrolase;  
 KW Lipoprotein.  
 FT CHAIN 2 69 COAT PROTEIN VP4.  
 FT CHAIN 70 323 COAT PROTEIN VP2.  
 FT CHAIN 324 565 COAT PROTEIN VP3.  
 FT CHAIN 566 862 COAT PROTEIN VP1.  
 FT CHAIN 863 1012 CORE PROTEIN P2A.  
 FT CHAIN 1013 1111 CORE PROTEIN P2B.  
 FT CHAIN 1112 1440 CORE PROTEIN P2C.  
 FT CHAIN 1441 1526 CORE PROTEIN P3A.  
 FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VP6.  
 FT CHAIN 1549 1731 PICORNAIN 3C.  
 FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.  
 FT LIPID 2 2 N-myristoyl glycine (in host) (By similarity).  
 FT ACT SITE 1695 1695 PROTEASE (POTENTIAL).  
 FT ACT SITE 1709 1709 PROTEASE (POTENTIAL).  
 FT SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;  
 Query Match 4.8%; Score 139.5; DB 1; Length 2193;  
 Best Local Similarity 20.3%; Pred. No. 0.056;  
 Matches 98; Conservative 57; Mismatches 141; Indels 187; Gaps 24;  
 QY 12 DGSTANLVV-----EVNN--EVMALEPVVGAIAAPVAGQNVDPWIRNNP- 56  
 DB 341 DGVSAPIUGFHPPTPIHIGEVHLLLEICRVETILEVNNL-----KNTETPMORLCPP 395  
 QY 57 --VQAPGGEFTVSPRNAPGEILWSAPLGFNLNLYSLHARMYNGVAGGFVQVILAGNAF 114  
 DB 396 VSVQSKTCELCAAFRADPGRD-----GPMQSTILGQLCRYVTQWSGSLVTFMFAGSPM 449  
 QY 115 TAGKIIFAAVPP--NPTTEGLSPSQVTWFFPHIIVDVRQLEPVLIPLDVRNNFYHNSQN 172

Db 450 ATGKMLIATTPGCVNPADRIATMLGT---HVWDFGLQSSVTLVVPWISNTHY----- 500  
QY 173 DSTIKLIAMLYPLRAN-NAGDDVFTVSCVRLTRSPDFDFELVPPVTSRTKPFVPI 231  
Db 501 -----RAHARAG-----YFDY-----YTTGI 516  
QY 232 LTVEEMSNRRFPLEKLYTGSSAFVQVQNGRCCTTGVLLGTQLTQLSAVNICFRGDTV 291  
Db 517 IITVQTNVVPFI-----GAPTTAYIVA----- 539  
QY 292 HIAGSHD-YTNMLASQNSNYDPTTEIPAPLGTDFVKGIOG-----MLTOTTRED----- 341  
Db 540 -LAAAQDNFTMKLCKD-----TEDI-----EQTANIQDPTIADMDIQTVNNQVRS 584  
QY 342 -----GSTRAHKATVSTGSHFTPKLGSVQVTTDNDNFOTGQNTKFTTPVGVITQD 391  
Db 585 LTAQLVPLTAADTEASSHRLGTGW---PALQAA-----ETGASSNASDKNLIET 631  
QY 392 G---NNHQEPQOWLPVNSGRTHGNVHLAPAVATFFGEOQLLFFRSTMPGCGSYPNNML 448  
Db 632 RCVLNHST--OETAIGNFFSRAG-----LVSITMTPTGTQ-----NTDGYVNWDI 676  
QY 449 DCL 451  
Db 677 DLM 679

RESULT 11  
POLG HE701  
ID -POLG HE701 STANDARD; PRT; 2194 AA.  
AC P32537;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Genome protein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.26) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
DE Human enterovirus 70 (strain J670/71) (EV 70).  
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.  
OC NCBI\_TaxID=31915;  
OX [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=91037960; PubMed=2172447;  
RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D., Minor P.D., Almond J.W.;  
RT "The complete nucleotide sequence of enterovirus type 70: relationships with other members of the picornaviridae.";  
RL J. Gen. Virol. 71:2291-2299(1990).  
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.  
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polypeptide. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {N} (N).  
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.  
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; D00820; BAA18891.1; -.

PIR: A36253; GNNYET.  
HGSP: P03300; IPOV.  
DR MEROPS; C03-UPA; -.  
DR INTERPRO: IPR003593; AAA\_ATPase.  
DR InterPro: IPR004004; Calici\_pol\_hel.  
DR InterPro: IPR009003; Cys\_Ser\_trypsin.  
DR InterPro: IPR001199; Pept\_3C\_picorn.  
DR InterPro: IPR000081; Peptidase\_C3.  
DR InterPro: IPR001138; Pico\_P1A.  
DR InterPro: IPR002527; Pico\_P2B.  
DR InterPro: IPR001676; Rhv.  
DR InterPro: IPR000605; RNA\_helicase.  
DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro: IPR001205; RNA\_pol\_P3D.  
DR InterPro: IPR007094; RNA\_pol\_PSVir.  
DR InterPro: IPR008975; Viral\_cap\_coat.  
DR Pfam: PF00548; Cys-protease\_3C; 1.  
DR Pfam: PF02226; Pico\_P1A; 1.  
DR Pfam: PF00947; Pico\_P2A; 1.  
DR Pfam: PF01552; Pico\_P2B; 1.  
DR Pfam: PF00073; rhv\_3.  
DR Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
DR PRINTS: PR00910; RNA\_helicase; 1.  
DR PRINTS: PR00918; CALICVIRUSNS.  
DR ProDom: PD001125; Cys\_protease\_3C; 1.  
DR ProDom: PD001306; Pico\_P2A; 1.  
DR ProDom: PD001274; Pico\_P2B; 1.  
DR SMART; SM00382; AAA\_1.  
KW Polyprotein; Coat protein; Core protein; Transferase; RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate; Lipoprotein.  
FT CHAIN 2 69 COAT PROTEIN VP4.  
FT CHAIN 70 319 COAT PROTEIN VP2.  
FT CHAIN 320 561 COAT PROTEIN VP3.  
FT CHAIN 562 871 COAT PROTEIN VP1.  
FT CHAIN 872 1014 CORE PROTEIN P2A.  
FT CHAIN 1015 1113 CORE PROTEIN P2B.  
FT CHAIN 1114 1443 CORE PROTEIN P2C.  
FT CHAIN 1444 1532 CORE PROTEIN P3A.  
FT CHAIN 1533 1554 GENOME-LINKED PROTEIN VP6.  
FT CHAIN 1555 1737 PICORNAIN 3C.  
FT CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.  
FT LIPID 2 N-myristoyl glycine (in host) (By similarity).  
FT ACT SITE 1701 PROTEASE (POTENTIAL).  
FT ACT\_SITE 1715 PROTEASE (POTENTIAL).  
FT SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;  
Query Match 4.6%; Score 133.5; DB 1; Length 2194;  
Best Local Similarity 23.7%; Pred. No. 0.16; Indels 47; Gaps 9;  
Matches 57; Conservative 34; Mismatches 103;  
QY 54 NNFVQAPGGE---FTVSPRNPAGFELWSAPL-----GPDLPYLSHLARMNGYAGGFV 105  
Db 375 NNVDASGVRLRVQISAGSDMDQLLEPLDLEGLPLRLTLGNISRYXTHWSSGLEM 434  
QY 106 QVILAGNAFTAGKIIFAAVPPNPFTEGLSPS---QVTMPFHIIIVDRQLEPVLILPDPVR 162  
Db 435 TFMFCGFSMTTKGLIICVTPPG---GSSPTDRMQAMLAATHVVDVDFGLQSSITIIIPWIS 490  
QY 163 NNFYHYNQNSDSTIK-----LIAMLYTPLRAN-NAGDDVFTVSCRV-----LTPSP 208  
Db 491 GSHYRMFTDAAKAINANVGVVTCFMQTNLVAPVGADQCYIVGMVAACKDFNLRMRDSP 550  
QY 209 DF-----DFIFLVPTVESRTKP---FTVPILTVEMSNRRFPIPLEKLY 250  
Db 551 DIQSAILEPQAATTQIGIVKTVANTVESEIKAEGLGVPSLNAVETGATSNTEPBEAIQ 610  
QY 251 T 251  
Db 611 T 611

```

RESULT 12
Y741_CHLMU STANDARD; PRT; 1007 AA.
AC Q9PJT6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein TC0741 precursor.
GN TC0741.
OS Chlamydia muridarum.
OC Bacteria; Chlamydiae; Chlamydiales; Chlamydiaceae; Chlamydia.
OX NCBI_TaxID=83560;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20150255; PubMed=10684935;
RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
RA Gwinn M., Nelson W., DeBoy R., Kolonay J., McClarty G., Salzberg S.L.,
RA Eisen J., Fraser C.M.;
RT "Genome sequences of Chlamydia trachomatis MoPn and Chlamydia
RT pneumoniae AR39."
RL Nucleic Acids Res. 28:1397-1406(2000).
CC -!- SIMILARITY: BELONGS TO THE CHLAMYDIAL CPN0572/CT456/TC0741
CC FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AE002342; AAF39550.1; -.
CC PIR; H81670; H81670.
CC TIGR; TC0741; -.
CC SIGNAL 1 51
CC CHAIN 52 1007 HYPOTHETICAL PROTEIN TC0741.
CC SEQUENCE 1007 AA; 104006 MW; 842800C0871B1518 CRC64;
CC
CC Query Match
CC Best Local Similarity 23.4%; Pred. No. 0.089;
CC Matches 75; Conservative 39; Mismatches 128; Indels 79; Gaps 14;
QY 219 TVESRTKPTVPIITVEEMNSRFPPIPLEKLYTGSSAFVQVQNGRCCTTGGVLLGTQL 278
Db 2 TTPISNPSPIPTVTVTSTTASSGSLGTSTVSTSTTSTVAQTAT---TTSSASTSIQS 58
QY 279 SAVNICTFRGDVTHIAGSHDVTWNLASQWNSYDTEETPAFLGTPDFVKGKQGLMTQTT 338
Db 59 SGENIQSTGTSPIT-----SSVSTAPSASATANKTSVAVGSKI---TSQBT 106
QY 339 RDGSGTRAHKAIVTSGSVHF-----TPKLGSG-----VQYTTDNDNDFQTGN 380
Db 107 SESET---QATSDGEVSSNYDDVDTPNNSDSTVDSYQDVETQYKTIENN-----GEN 159
QY 381 TKFTPVGVIGQNNHQNHPQWLVNYSRGTGNVH-----LAPAVAPTFPGQLLF 432
Db 160 T-YETIGSHGKXNTHQVE-----SHASGTGNPINNQEAIRQLRSSTYTTSPRNEIF 211
QY 433 FRSTWPGSGYNNMLDCLLPQEWVQHFQCEAPQAQSDVALLRFV-NPDTGRVLFECKLH 491
Db 212 ----SPGPEGLNMLSPSPSPD-----KSSLIAFLSNPNNTAKMLE----- 249
QY 492 KSGYVTVTAHTGPHDLVIPNG 512
Db 250 HSHGLVFDITTRSSFIYPNG 270

```

```

RESULT 13
POLG_CX16G STANDARD; PRT; 2193 AA.
AC Q65900;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
DE linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Coxsackievirus A16 (strain G-10).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=69159;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94303216; PubMed=8030260;
RA Poyry T., Hyypia T., Horsnell C., Kinnunen L., Hovi T., Stanway G.;
RA "Molecular analysis of coxsackievirus A16 reveals a new genetic group
RT of enteroviruses."
RL Virology 202:982-987(1994).
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular
CC membranes and is associated with viral RNA synthesis.
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA] (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U05876; AAA50478.1; -.
CC HSPF; P03300; IPOV.
CC MEROPS; C03.022; -.
CC InterPro; IPR003593; AAA ATPase.
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000399; Pept_3C_picorn.
CC InterPro; IPR000081; Peptidase_C3.
CC InterPro; IPR003138; Pico_P1A.
CC InterPro; IPR002527; Pico_P2B.
CC InterPro; IPR01676; Rhv.
CC InterPro; IPR006005; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_Psvir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00548; Cys-protease-3C; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF00073; rhv; 3.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICVIRUSNS.
CC ProDom; PD001125; Cys_protease_3C; 1.
CC ProDom; PD001306; Pico_P2A; 1.

```

DR ProDom; PD001274; Pico\_P2B; 1.  
DR SMART; SM00382; AAA; 1.  
KW Polyprotein; Coat protein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
KW Lipoprotein. 2 69 COAT PROTEIN VP4.  
FT CHAIN 70 323 COAT PROTEIN VP2.  
FT CHAIN 324 565 COAT PROTEIN VP3.  
FT CHAIN 566 862 COAT PROTEIN VP1.  
FT CHAIN 863 1012 CORE PROTEIN P2A.  
FT CHAIN 1013 1111 CORE PROTEIN P2B.  
FT CHAIN 1112 1440 CORE PROTEIN P2C.  
FT CHAIN 1441 1526 CORE PROTEIN P3A.  
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VPG.  
FT CHAIN 1549 1731 PICORNAIN 3C.  
FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.  
FT LIPID 2 2 N-myristoyl glycine (in host) (By similarity)  
FT ACT SITE 1695 1695 PROTEASE (POTENTIAL).  
FT ACT SITE 1709 1709 PROTEASE (POTENTIAL).  
SQ SEQUENCE 2193 AA; 243209 MW; 04B3BCE572A76E38 CRC64;  
  
Query Match 4.4%; Score 127.5; DB 1; Length 2193;  
Best Local Similarity 20.4%; Pred. No. 0.44;  
Matches 98; Conservative 52; Mismatches 149; Indels 181; Gaps 24;  
  
QY 12 DGSANLVP-----EVNN--EVMALEPVVGAIAAPVACQNVDPWIRNPF- 56  
DB 341 DGVSAPIPLGFHTPAIHIPGEVNNLEICRVETILEVNNL-----QSNETTPMORLCFP 395  
QY 57 --VOAGGFTVSPNAPGELLWSAPILGDLNPLSLHARMYNGVAGGVEVOVILAGNAF 114  
DB 396 VSVQSKTGELCAVFRADPGR-----NGPQSTILGQLCRYTQWGSLEVTMPAGSEF 449  
QY 115 TAGKIIFAAVPP--NPTTEGLSPQVTMFPHIIVDRQLEPVLPLPVDVNNFYHYNQSN 172  
DB 450 ATGKMLIAYTPPGGVPAADRLTAMLTG---HVIWDFGLQSSVTLVPIWISNTHY----- 500  
QY 173 DSTIKLIAMLYPLRANNAGDVFVTSRVLTRPSDFDFILVPTVESRTKPFVPIVL 232  
DB 501 -----RA-HAKDGY-----FDY-----YITGTI 517  
QY 233 TVEEMSNRRPIPLEKLYTGPSSAFVVPQNGCRCTTDDGVLLGTTQLSAVNICTFRGDVTH 292  
DB 518 TIWYQTNVVPV-----GAPTAIVV----- 539  
QY 293 IAGSHD-YTMNLASQNSNDPTBEIPAPL-----GTPDFV-----GKIQGMLT--QTTR 339  
DB 540 LAAQDNFTMKLCKD-----TEDIEQSANIQDGIADMDIDQAVTSRVRGRLTSLQVEP 592  
QY 340 EDGSTRAHKATVSTGSHFTPKLGSVQYTTDTNNDFOGTQNTKFTPGVVIQDGN----- 393  
DB 593 TAANTNASEHRLGTGLV---PALQAAETGASSN-----AQDENLLETTC 633  
QY 394 --NHNEPQQWLVNYSRGTCHNVHLPAPAVAPTPEGQQLLFFRSTMPGCGSGYNNMLDCL 451  
DB 634 VLNH-HSTQETIGNFFSRAG---LVSIIIMPITGTQ-----NTDGVYVNDIDLM 679  
  
RESULT 14  
POLG FMDVO  
ID POLG FMDVO STANDARD; PRT; 2332 AA.  
AC P03305;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Genome polyprotein [Contains: Nonstructural protein P2A; Coat  
DE protein VP4; Coat protein VP2; Coat protein VP3; Coat protein VP1;  
DE Core protein P12; Core protein P14; Core protein P14; Genome-linked  
DE protein VPG; Protease (EC 3.4.22.-); RNA-directed RNA polymerase  
DE (EC 2.7.7.48)].  
OS Foot-and-mouth disease virus (strain O1) (FMDV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Aphthovirus.  
OX NCBI\_TaxID=73482;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O1K;  
RX MEDLINE=84297249; PubMed=6089122;  
RA Forss S., Strebel K., Beck E., Schaller H.;  
RT "Nucleotide sequence and genome organization of foot-and-mouth  
RL disease virus.";  
RN Nucleic Acids Res. 12:6587-6601(1984).  
[2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O1BFS;  
RX MEDLINE=83143292; PubMed=6298715;  
RA Makoff A.J., Paynter C.A., Rowlands D.J., Boothroyd J.C.;  
RT "Comparison of the amino acid sequence of the major immunogen from  
RT three serotypes of foot and mouth disease virus.";  
RL Nucleic Acids Res. 10:8285-8295(1982).  
[3]  
RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
RX MEDLINE=89143740; PubMed=2537470;  
RA Acharya R., Fry E., Stuart D., Fox G., Rowlands D., Brown F.;  
RT "The three-dimensional structure of foot-and-mouth disease virus at  
RT 2.9-A resolution.";  
RL Nature 337:709-716(1989).  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA}(N).  
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
CC each of which is composed of one copy each of proteins VP1, VP2,  
CC VP3, and VP4.  
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC -!- MISCELLANEOUS: THE STRAIN O1K SEQUENCE IS SHOWN. THERE ARE NO RESTRICTIONS ON ITS  
CC -!- MISCELLANEOUS: THE COAT PROTEIN VP1 CONTAINS THE MAIN ANTIGENIC  
CC DETERMINANTS OF THE VIRION; THEREFORE, CHANGES IN ITS SEQUENCE  
CC MUST BE RESPONSIBLE FOR THE HIGH ANTIGENIC VARIABILITY OF THE  
CC VIRUS.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X00871; CAA25416.1; -;  
CC EMBL; J02185; AAB42635.1; -;  
CC FDB; IQM1; 18-SEP-01.  
CC FDB; IQOL; 10-NOV-00.  
CC MEROPS; C03.008; -;  
CC InterPro; IPR004004; Calici\_pol\_hel.  
CC InterPro; IPR009003; Cys\_ser\_trypsin.  
CC InterPro; IPR004080; FMDV\_Picoat.  
CC InterPro; IPR008739; Peptidase\_C28.  
CC InterPro; IPR001676; Rnv.  
CC InterPro; IPR000605; RNA\_helicase.  
CC InterPro; IPR007095; RNA\_pol\_DS\_PS.  
CC InterPro; IPR001205; RNA\_pol\_P3D.  
CC InterPro; IPR007094; RNA\_pol\_PSVir.  
CC InterPro; IPR008975; Viral\_cap\_coat.  
CC Pfam; PF05408; Peptidase\_C28; 1.  
CC Pfam; PF00073; Rnv; 3.  
CC Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
CC Pfam; PF00910; RNA\_helicase; 1.  
CC PRINTS; PR00918; CALICIVIRUSNS.  
CC PRINTS; PR01542; FMDV\_PICOAT.  
CC Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;  
KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;  
KW Myristate; 3D-structure; Lipoprotein.  
FT CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.  
FT CHAIN 202 286 COAT PROTEIN VP4.  
FT CHAIN 287 504 COAT PROTEIN VP2.  
FT FT



FT CHAIN 505 724 COAT PROTEIN VP3.  
 FT CHAIN 725 937 COAT PROTEIN VP1.  
 FT CHAIN 938 1307 CORE PROTEIN P12.  
 FT CHAIN 1108 1425 CORE PROTEIN P34.  
 FT CHAIN 1426 1578 CORE PROTEIN P14.  
 FT CHAIN 1579 1649 GENOME-LINKED PROTEIN VPG.  
 FT CHAIN 1650 1862 PROTEASE.  
 FT CHAIN 1863 2332 RNA-DIRECTED RNA POLYMERASE.  
 FT LIPID 202 202 N-myristoyl glycine (in host).  
 FT DISULFID 511 511 INTERCHAIN (IN VP3 DIMER).  
 FT VARIANT 780 780 I -> V (IN STRAIN O1BFS).  
 FT VARIANT 808 808 G -> R (IN STRAIN O1BFS).  
 FT VARIANT 861 861 N -> S (IN STRAIN O1BFS).  
 SQ SEQUENCE 2332 AA; 258924 MW; 4A83176F43447D68 CRC64;

Query Match 4.4%; Score 127; DB 1; Length 2332;  
 Best Local Similarity 22.3%; Pred. No. 0.52;  
 Matches 73; Conservative 39; Mismatches 134; Indels 82; Gaps 14;

QY 92 LARMYNGYAGGEVQVILAGNAFTAGKIIFAAVPPNFTPEGLSPQVMTFPHIIVDVROL 151  
 DB 380 LTDSYAYMNGWDVEVTAQVNGFNCGCLLVAMVPELYSQKRELYQLTLFPHQFINPRN 439  
 QY 152 EPLVPLDPRNRYFY-HYNQSDSTIKLIAMLYTHLRANNAGDDVFTVSCRVLTRPSDF 210  
 DB 440 MTAHITVPFVGVRNDQYKVRPWT--LVVMVAPLTVNTEG-----APQI 483  
 QY 211 DRIFLVPPVTSRTPFTVPILTVEMSNRRPIPLEKLY-----TGPSSAFVW----- 259  
 DB 484 KYANIAPT-----NVHAGEPPSKEGIFPVACSDGYGGLVTTDPKTADPVYGVKN 535  
 QY 260 QPON---GRCTTGVLLGTTQLSAVNICTFRGDVTHIAGSHD-----YTNLASQNMW 309  
 DB 536 PPRNQLPGRFTN---LLDVAE-ACPTFLFEGGVPTVTKTSDRVLQAFDMSLAQKMS 591  
 QY 310 NYDPTTEIPAPLPTDPTDFVKIQGLMLTQTTREDGSTRAHKATVSTG-----SVHFTPK 361  
 DB 592 N-----TFLAGLAQYTYQVSGTINLH--FMFTGTDAKARYMVAYAPP 632  
 QY 362 LGSVQVYTTDT-----NNDFQTGONTKFT 384  
 DB 633 GMEPKPTPEAAAHCHIAEWDGTGLNSKFT 660

## RESULT 15

POLG\_FMDV1  
 ID POLG\_FMDV1 STANDARD; PRT; 2333 AA.  
 AC P03306; Q64768; Q84750; Q84751; Q84752; Q84753; Q84754; Q84760;  
 AC Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;  
 AC Q84769; Q89824;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat  
 proteins VP1 TO VP4; Core protein P52; Genome-linked proteins VPG1 TO  
 DE VPG3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed  
 DE RNA polymerase P56A (EC 2.7.7.48)].  
 OS Foot-and-mouth disease virus (strain A10-61) (Aphthovirus A) (FMDV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Aphthovirus.  
 OK NCBI\_TaxID=12112;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84169547; PubMed=6324120;  
 RA Carroll A.R., Rowlands D.J., Clarke B.E.;  
 RT "The complete nucleotide sequence of the RNA coding for the primary  
 translation product of foot and mouth disease virus.";  
 RL Nucleic Acids Res. 12:2461-2472 (1984).  
 RN [2]  
 RP SEQUENCE OF 115-1048 FROM N.A.  
 RX MEDLINE=82211814; PubMed=6282711;  
 RA Boothroyd J.C., Harris T.J.R., Rowlands D.J., Lowe P.A.;

"The nucleotide sequence of cDNA coding for the structural proteins  
 of foot-and-mouth disease virus.";  
 Gene 17:153-161 (1982).  
 CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the  
 poliovirus polyprotein. In other picornavirus reactions Glu may be  
 substituted for Gln, and Ser or Thr for Gly.  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 {RNA} (N).  
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 each of which is composed of one copy each of proteins VP1, VP2,  
 VP3, and VP4.  
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; V01130; CAA24361.1; -;  
 CC EMBL; X00429; CAA25127.1; -;  
 CC MEROPS; C03.008; -;  
 CC MEROPS; C28.001; -;  
 CC InterPro; IPR004004; Calici\_pol\_hel.  
 CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
 CC InterPro; IPR004080; FMDV\_P1coat.  
 CC InterPro; IPR008739; Peptidase\_C28.  
 CC InterPro; IPR001676; Rhv.  
 CC InterPro; IPR000605; RNA\_helicase.  
 CC InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 CC InterPro; IPR001205; RNA\_pol\_P3D.  
 CC InterPro; IPR007094; RNA\_pol\_P5vir.  
 CC InterPro; IPR008975; Viral\_cap\_coat.  
 CC Pfam; PF05408; Peptidase\_C28; 1.  
 CC Pfam; PF00073; rhv; 3.  
 CC Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 CC Pfam; PF00910; RNA\_helicase; 1.  
 CC PRINTS; PR00918; CALICIVIRUSNS.  
 CC PRINTS; PR01542; FMDV\_P1COAT.  
 KW Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;  
 KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;  
 KW Myristate; Lipoprotein.  
 FT CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.  
 FT CHAIN 202 286 COAT PROTEIN VP4.  
 FT CHAIN 287 504 COAT PROTEIN VP2.  
 FT CHAIN 505 725 COAT PROTEIN VP1.  
 FT CHAIN 726 937 COAT PROTEIN VP1.  
 FT CHAIN 938 1578 CORE PROTEIN P52.  
 FT CHAIN 1579 1601 GENOME-LINKED PROTEIN VPG1.  
 FT CHAIN 1602 1625 GENOME-LINKED PROTEIN VPG2.  
 FT CHAIN 1626 1649 GENOME-LINKED PROTEIN VPG3.  
 FT CHAIN 1650 1863 PROTEASE P20B.  
 FT CHAIN 1864 2333 RNA-DIRECTED RNA POLYMERASE P56A.  
 FT LIPID 202 202 N-myristoyl glycine (in host).  
 FT CONFLICT 396 396 S -> C (IN REF. 2).  
 FT CONFLICT 632 632 P -> L (IN REF. 2).  
 FT CONFLICT 632 632 P -> L (IN REF. 2).  
 SQ SEQUENCE 2333 AA; 259645 MW; 4FC667DCC521BC60 CRC64;

Query Match 4.4%; Score 126.5; DB 1; Length 2333;  
 Best Local Similarity 22.2%; Pred. No. 0.56;  
 Matches 78; Conservative 41; Mismatches 144; Indels 89; Gaps 15;

QY 72 PGEILWAPLPGDLPNYPYLSHARMYNGVAGGEVQVILAGNAFTAGKIIFAAVPPNFPTE 131  
 DB 360 PFGYITKLELPTDHRGVFGLVDSYAYMRNGDVEVSAGVNGQFGLLVAMVPEWKA 419  
 QY 132 GLSPSQVTMFPHIIVDVRLQLEPVLPLDPRNRYFY-HYNQSDSTIKLIAMLYTPLRAN 190  
 DB 420 TREKYQLTLFPHQFISPRNTNMTAHTVTPYLGVRNRDQYKHKPWT--LVVMVLSPLTVSN 477

Wed Jun 2 09:13:30 2004

```
QY 191 AGDDVFTVSCRVLTRPSDFDFIFLVPTV-----ESRTKPTVPILTVEEMNSRPPIP 245
Db 478 TA-----APQIKVYANIAPTYVHVAGELPSKEGIFPVACADGYGGLVTTDP 523
QY 246 -----LEKLYTGPSSAFVQPONGRCCTDGVLLGTTTOLSAVNIC-TF----- 286
Db 524 KTADPVYGVYNPPKTN-----PGRF-----TNLLDVAEACPTFLRFDGKPYVVT 570
QY 287 RGDVTHIAGSHDYTMNLASONMSNYDPTTEEIPAPLGTPDFVGKIQGMLTQTTREDGSTR 346
Db 571 RADDTRLLAKFD--VSLAAKMSN-----TYLSGIAQYVYQYSGTINL 611
QY 347 H-KATVSTGS-----VHFTPKLGSVQYTTDT-----NNDFOGQNTKFT 384
Db 612 HFMTGSTDSKARYMVAYIPP--GVETPPDTPEAAAHCIHAEDTGLNSKFT 661
```

Search completed: June 1, 2004, 13:47:16  
Job time : 8.51141 secs

Result No.	Score	Query Match	Length	DB	ID	Description
1	2804	97.2	539	2	S40111	capsid protein - h
2	1867.5	64.7	542	2	S60616	capsid protein - h
3	1161.5	40.2	530	2	B37471	capsid protein - N
4	1110	38.5	546	2	B37491	major capsid prote
5	337.5	11.7	576	2	A53982	capsid protein - E
6	329	11.4	2344	2	S64740	genome polyprotein
7	323	11.2	2344	2	RRWRH	genome polyprotein
8	319	11.1	2344	2	S55339	genome polyprotein
9	300.5	10.4	671	1	VCWVF9	coat protein - fel
10	296.5	10.3	702	1	A48562	coat protein - San
11	294	10.2	668	1	VCWVFC	coat protein - fel
12	293.5	10.2	668	2	QJ2354	capsid protein - f
13	293	10.2	668	1	VCWVFF	coat protein - fel
14	288.5	10.0	703	1	A48562	coat protein - San
15	288	10.0	668	2	QJ2356	capsid protein - f
16	133.5	4.6	2194	1	GNNYE7	genome polyprotein
17	130.5	4.5	1007	2	H81670	conserved hypothet
18	127	4.4	2332	1	GNNYF	genome polyprotein
19	126.5	4.4	2333	1	GNNY2F	genome polyprotein
20	126	4.4	948	2	T11678	hypothetical prote
21	123	4.3	26926	1	I38344	titin, cardiac mus
22	122.5	4.2	940	2	D89723	protein F39D8.1b [
23	122.5	4.2	945	2	T21998	hypothetical prote
24	122.5	4.2	2201	1	GNNYA9	genome polyprotein
25	122.5	4.2	1011	1	GNNYC1	genome polyprotein
26	121.5	4.2	733	2	QJ1892	capsid protein - f
27	121	4.2	3473	1	A46112	genome polyprotein
28	121	4.2	3473	2	S27927	polyprotein - rice
29	120.5	4.2	70183	2	QJ1891	capsid protein - f

Db 421 VAPTFPGQLFFRSTMPGCGYPNNMLDCLLPQEWLHFYQEAAPQAQSDVALLRFVNP 480  
QY 481 TGRVLPCKLHKSGVYVTAHTGPHDLVPPNGYFRFDSWVNOFYTLAPMGAGRRAL 539  
Db 481 TGRVLPCKLHKSGVYVTAHTGPHDLVPPNGYFRFDSWVNOFYTLAPMGAGRRAL 539

RESULT 2  
S60616  
A:Title: capsid protein - human calicivirus (strain Melksham)  
C:Species: human calicivirus  
A:Variety: strain Melksham  
C:Date: 23-May-1997 #sequence\_revision 23-May-1997 #text\_change 28-Jul-2000  
C:Accession: S60616  
R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
Virus Res. 37, 271-283, 1995  
A:Title: Capsid diversity in small round-structured viruses: molecular characterization  
A:Reference number: S60615; MUID:96136658; PMID:8533462  
A:Accession: S60616  
A:Molecule type: Genomic RNA  
A:Residues: 1-542 <GRE>  
A:Cross-references: EMBL:X81879; NID:g976077; PIDN:CAA57462.1; PID:g976079  
A:Experimental source: Strain Melksham  
A:Note: it is uncertain whether Met-1 or Met-3 is the initiator  
C:Superfamily: human calicivirus capsid protein  
C:Keywords: capsid protein; coat protein

Query Match 64.7%; Score 1867.5; DB 2; Length 542;  
Best Local Similarity 63.3%; Pred. No. 1.6e-130;  
Matches 345; Conservative 79; Mismatches 108; Indels 13; Gaps 4;  
QY 1 MKMASNDANPSDGTANLVEPVNNEVMALEPVVGAIAAPVAGQOQNVDPWIRNPFVQAP 60  
Db 1 MKMASNDAASTDGAAGLVNESNNEVMALEPVVGAIAAPVGTQTIIDPWIRANFVQAP 60

QY 61 GGEFTVSPNAPGEILWSAPLGPDLNPLYSLHARMYNGVAGGEVQVILAGNAFTAGKII 120  
Db 61 NGEFTVSPNAPGEILNLSLGPDLNPLYSLHARMYNGVAGGEVQVILAGNAFTAGKLV 120  
QY 121 FAAPVPNFPTEGLSPQVTFPHIIVDVROLEPVLIPDPVRNFFHYNQSDNSTIKLIA 180  
Db 121 FAAPVPNFPTEGLSPQVTFPHIIVDVROLEPVLIPDPVRNFFHYNQSDNSTIKLIA 180

QY 181 MLYTPLRANAGDDVFTVSCRVLTRSPDFDFLVPPTVESRTKPTFPIITVEEMSNS 240  
Db 181 MLYTPLRANAGDDVFTVSCRVLTRSPDFDFLVPPTVESRTKPTFPIITVEEMSNS 240  
QY 241 RFPILPEKLYTGSSAFVQPNQGRCTTDGVLGTTQLSARNICTFRGDVTHIAGSHDYT 300  
Db 241 RFPVPIIDMTYSPNEVSVQCNQGRCTTDGVLGTTQLSARNICTFRGDVTHIAGSHDYT 300

QY 301 MN--LASQNSNYDPTBEIPAPLGTDFVGKIQGLMTQTTRDQ-----STRAHKATVS 352  
Db 301 NNVTITNLGSPDFSEDIAPLGPVDFQGRVFGVISORDKQNAAGHSEPANRGHDAVVP 360  
QY 353 TGSVHFTPKLGSVQYTTDNDFTQNTKFTPGVIGVQDGNHNEPQOQVLPVNSGRTG 412  
Db 361 TYTAQYTPKLQIQIGTWTQTDLTVNQVFKFTPVGL--NDTEHFN---QWVVPRIYAGLN 415

QY 413 HNVHLAPAVAPTFPGEQLFFRSTMPGCGYPNNMLDCLLPQEWVQHFQEAAPQAQSDVA 472  
Db 416 LNTNLAPSVAVPFGEERLLFRSHLP.LKGGYGNFAIDCLLPQEWVQHFQEAAPSMSEVA 475  
QY 473 LLRFVNPDTGRVLPCKLHKSGVYVTAHTGPHDLVIPPNGYFRFDSWVNOFYTLAPMGNG 532  
Db 476 LVRYINPDTGRALFEAKLHRAGFTVSSNTSAPVVVPANGYFRFDSWVNOFYSLAPMGTG 535

QY 533 AGRRR 537  
Db 536 NGRRR 540

## RESULT 3

B37471  
A:Title: capsid protein - Norwalk virus  
C:Species: Norwalk virus  
C:Date: 24-Feb-1994 #sequence\_revision 25-Apr-1997 #text\_change 28-Jul-2000  
C:Accession: B37471  
R:Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.  
Virology 195, 51-61, 1993  
A:Title: Sequence and genomic organization of Norwalk virus.  
A:Reference number: A37471; MUID:93303939; PMID:8391187  
A:Accession: B37471  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: genomic RNA  
A:Residues: 1-530 <JIA>  
A:Cross-references: GB:M87661; NID:g1061311; PIDN:AAB50466.1; PID:g1061313  
A:Note: sequence extracted from NCBI backbone (NCBI:P:134157)  
C:Superfamily: human calicivirus capsid protein

Query Match 40.2%; Score 1161.5; DB 2; Length 530;  
Best Local Similarity 45.4%; Pred. No. 3.1e-78;  
Matches 250; Conservative 83; Mismatches 181; Indels 37; Gaps 14;  
QY 1 MKMASNDANPS-DGST--ANLVEPVN-NEVMALEPVVGAIAAPVAGQOQNVDPWIRNPF 56  
Db 1 MMASKDATSVDCASGAGQLVPEVNASDFLAMPVAGSTAVATAGVNPIDPWIINNF 60

QY 57 VOAPGGEFTVSPNAPGEILWSAPLGPDLNPLYSLHARMYNGVAGGEVQVILAGNAFTA 116  
Db 61 VOAPGGEFTVSPNAPGEILWSAPLGPDLNPLYSLHARMYNGVAGGEVQVILAGNAFTA 120  
QY 117 GKIIIPAAVPPNFPTEGLSPQVTFPHIIVDVROLEPVLIPDPVRNFFHYNQSDNSTI 176  
Db 121 GKIIIPAAVPPNFPTEGLSPQVTFPHIIVDVROLEPVLIPDPVRNFFHYNQSDNSTI 180

QY 177 KLIAMLYTPLRANAGDDVFTVSCRVLTRSPDFDFLVPPTVESRTKPTFPIITVEE 236  
Db 181 RLVCMLYTPLRANAGDDVFTVSCRVLTRSPDFDFLVPPTVESRTKPTFPIITVEE 240  
QY 237 MSNRFPTPEKLYTGSSAFVQPNQGRCTTDGVLGTTQLSARNICTFRGDVTHIAGS 296  
Db 241 LNSRAPLPISSIGISPDNVQSVQFQNGRCTLGRLVGTTPVLSLHVAKIRGT-----S 294

QY 297 HDYTNLASQNSNYDPTBEIPAPLGTDFVGKIQGLMTQTTRDGGSTRAHKATVSTGVS 356  
Db 295 NGTVNLINTELDTGTPHPFEG-PAPITGPD-LGGCDWHINMT--QFGHSSQYQYDVTTPD 350  
QY 357 HFTPKLGSVQYTTDNDFTQNTKFTPGVIGQ--DGNHNEPQ--QWVLPVNSGRTG 412  
Db 351 TEVPHLGSIQ-----ANGIGSNY-----VGVLWSISPSPHSPSQVDLWKIPNYGSSIT 400

QY 413 HNVHLAPAVAPTFPGEQLFFRSTMPGCGYPNNMLDCLLPQEWVQHFQEAAPQAQSDVA 472  
Db 401 EATHLAPSVYPPGFEVLVFFEMSKMPGAY---NLCLLPQEVYISHLASEQAPTVGBAA 457  
QY 473 LLRFVNPDTGRVLPCKLHKSGVYV---AHTGPHDLVIPPNGYFRFDSWVNOFYTLAP 528  
Db 458 LLHYVDPDTGRNLGFKAYPDGFLCTVPNGASSGSQL--PINGVVFVWSVRSFYQLKP 515

QY 529 MCGAGRRAL 539  
Db 516 VGTASSARGRL 526

## RESULT 4

B37491  
A:Title: capsid protein [similarity] - Southampton virus  
N:Alternate names: orf2 protein  
C:Species: Southampton virus  
C:Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 18-Aug-2000  
C:Accession: B37491  
R:Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
Science 259, 516-519, 1993  
A:Title: Sequence and genome organization of a human small round-structured (Norwalk-like)



C;Keywords: blocked amino end; polyprotein

Query Match 11.4%; Score 329; DB 2; Length 2344;  
Best Local Similarity 25.3%; Pred. No. 1.3e-15;  
Matches 137; Conservative 62; Mismatches 213; Indels 130; Gaps 23;

QY 4 ASDNANPSDGTANLVPEVNNVMALEPVVGAIAAPVAGQONVID---PWIRNNFVOAP 60  
DB 1785 ASVPGTTTDDGMDPGVV--ATTSVTAENSSASIAATAGIGGPPQVDDQETW-RTNFYY-- 1839

QY 61 GGEFTVSPRNAPGEILMSAPLGPDLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120  
DB 1840 NDVFTWSVADAPGSILYTVQHSFQNNPFTAVLSQMYAGWAGGQFRFIVAGSGVFGRLV 1899

QY 121 FAVPNNFTEGLSPS-QVTMFPPIIIVDVRQLEPVLPLPDVRRNFYHYNQSNDSITIKLI 179  
DB 1900 AAVIPPGIE---IGPGLVRQFPFHVVIDARSLSEPVTITMPDLRPNMYHTPGDGLVPTLV 1956

QY 180 AMLYTLRANAGDDVFTVSCRVLTRPSDFDFIFLVPV---TVESRTKP--FTVPILTV 234  
DB 1957 LSVYNNL-INPFGGTSAIQVTVETRPSEDFEFWMIRAPSSKTVDISIPAGLLTTPVLT- 2014

QY 235 EEMSNRRFPILEKLYTGPSSAFVQPPQNG---RCTTDGVLGTTQ-LSAVNICTRGDV 290  
DB 2015 -----GVGNDNRWG--QIVGLQVPVGGFSTCNRHWNLSGTWSSPRF---GDI 2060

QY 291 THIAGSHDYTMNLAS---QNW-----SNYD--PTEETPAPLGTDP----- 325  
DB 2061 DHRGSASYPGNATNVLFQWYANAGSAIDNPISQV-APDGFDMSPVFPNGFGIPAAGW 2119

QY 326 -----FVGKIQGMLTOTTRDGGSTRAHKA---TVSTGS--- 355  
DB 2120 VGFCAIWNNSGAPNVTTVQAYELGFATGAPGNLQPTTNTSGAQTVAKSIYAVVTGTAQN 2179

QY 356 -----VHFTPKLGSVQYTTDTNNDFTQGTQKTFPVGVIIQ----- 390  
DB 2180 PAGLFVWASGIITSPNAQRIITYTPQDRIIVTTPGTAAAPVGKNTPIMFASVVRRTGDVN 2239

QY 391 -----DGNHNEPQQWVLPNYSRGTGNVHLAPAVPTFGEQLLFFRSTMPGCGYPNM 446  
DB 2240 ATAGSANGTQGTGSQLPVTIGLSLNNYSSA-----LMPGQ---FFVWQLTFASGFMEI 2291

QY 447 NL 448  
DB 2292 GL 2293

RESULT 8

S55399  
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)  
C;Species: rabbit hemorrhagic disease virus  
A;Variety: isolate BS89  
C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 18-Jun-1999  
C;Accession: S55399  
R;Rossi, C.  
submitted to the EMBL Data Library, May 1995  
A;Reference number: S55399  
A;Accession: S55399  
A;Molecule type: genomic RNA  
A;Residues: 1-2344 <ROS>  
A;Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641  
A;Experimental source: isolate BS89  
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C;Keywords: polyprotein

Query Match 11.1%; Score 319; DB 2; Length 2344;  
Best Local Similarity 25.1%; Pred. No. 7e-15;  
Matches 136; Conservative 64; Mismatches 213; Indels 128; Gaps 24;

QY 4 ASDNANPSDGTANLVPEVNNVMALEPVVGAIAAPVAGQONVID---PWIRNNFVOAP 60  
DB 1785 ASVPGTTTDDGMDPGVV--ATTSVTAENSSASIAATAGIGGPPQVDDQETW-RTNFYY-- 1839

QY 61 GGEFTVSPRNAPGEILMSAPLGPDLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120  
DB 1840 NDVFTWSVADAPGSILYTVQHSFQNNPFTAVLSQMYAGWAGGQFRFIVAGSGVFGRLV 1899

QY 121 FAVPNNFTEGLSPS-QVTMFPPIIIVDVRQLEPVLPLPDVRRNFYHYNQSNDSITIKLI 179  
DB 1900 AAVIPPGIE---IGPGLVRQFPFHVVIDARSLSEPVTITMPDLRPNMYHTPGDGLVPTLV 1956

QY 180 AMLYTLRANAGDDVFTVSCRVLTRPSDFDFIFLVPV---TVESRTKP--FTVPILTV 234

C;Keywords: blocked amino end; polyprotein

Query Match 11.4%; Score 329; DB 2; Length 2344;  
Best Local Similarity 25.3%; Pred. No. 1.3e-15;  
Matches 137; Conservative 62; Mismatches 213; Indels 130; Gaps 23;

QY 4 ASDNANPSDGTANLVPEVNNVMALEPVVGAIAAPVAGQONVID---PWIRNNFVOAP 60  
DB 1785 ASVPGTTTDDGMDPGVV--ATTSVTAENSSASIAATAGIGGPPQVDDQETW-RTNFYY-- 1839

QY 61 GGEFTVSPRNAPGEILMSAPLGPDLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120  
DB 1840 NDVFTWSVADAPGSILYTVQHSFQNNPFTAVLSQMYAGWAGGQFRFIVAGSGVFGRLV 1899

QY 121 FAVPNNFTEGLSPS-QVTMFPPIIIVDVRQLEPVLPLPDVRRNFYHYNQSNDSITIKLI 179  
DB 1900 AAVIPPGIE---IGPGLVRQFPFHVVIDARSLSEPVTITMPDLRPNMYHTPGDGLVPTLV 1956

QY 180 AMLYTLRANAGDDVFTVSCRVLTRPSDFDFIFLVPV---TVESRTKP--FTVPILTV 234  
DB 1957 LSVYNNL-INPFGGTSAIQVTVETRPSEDFEFWMIRAPSSKTVDISIPAGLLTTPVLT- 2014

QY 235 EEMSNRRFPILEKLYTGPSSAFVQPPQNG---RCTTDGVLGTTQ-LSAVNICTRGDV 290  
DB 2015 -----GVGNDNRWG--QIVGLQVPVGGFSTCNRHWNLSGTWSSPRF---GDI 2060

QY 291 THIAGSHDYTMNLAS---QNW-----SNYD--PTEETPAPLGTDP----- 325  
DB 2061 DHRGSASYPGNATNVLFQWYANAGSAIDNPISQV-APDGFDMSPVFPNGFGIPAAGW 2119

QY 326 -----FVGKIQGMLTOTTRDGGSTRAHKA---TVSTGS--- 348  
DB 2120 VGFCAIWNNSGAPNVTTVQAYELGFATGAPGNLQPTTNTSGAQTVAKSIYAVVTGTAQN 2179

QY 349 -----ATVSTGSHVHTPKLGSVQYTTDTNNDFTQGTQKTFPVGVIIQ----- 390  
DB 2180 PAGLFVWASGIITSPNAQRIITYTPQDRIIVTTPGTAAAPVGKNTPIMFASVVRRTGDVN 2239

QY 391 -----DGNHNEPQQWVLPNYSRGTGNVHLAPAVPTFGEQLLFFRSTMPGCGYPNM 446  
DB 2240 ATAGSANGTQGTGSQLPVTIGLSLNNYSSA-----LMPGQ---FFVWQLTFASGFMEI 2291

QY 447 NL 448  
DB 2292 GL 2293

RESULT 7

R;WWRH  
genome polyprotein - rabbit hemorrhagic disease virus  
N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C;Species: rabbit hemorrhagic disease virus  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998  
C;Accession: A41039  
R;Meyer, G.; Wirblich, C.; Thiel, H.J.  
Virology 184, 664-675, 1991  
A;Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing  
A;Reference number: A41039; MUID:91361557; PMID:1840711  
A;Accession: A41039  
A;Molecule type: Genomic RNA  
A;Residues: 1-2344 <MEY>  
A;Cross-references: GB:M67473  
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C;Keywords: nucleotidyltransferase

Query Match 11.2%; Score 323; DB 1; Length 2344;  
Best Local Similarity 25.1%; Pred. No. 3.6e-15;  
Matches 136; Conservative 66; Mismatches 210; Indels 130; Gaps 24;

QY 4 ASDNANPSDGTANLVPEVNNVMALEPVVGAIAAPVAGQONVID---PWIRNNFVOAP 60  
DB 1785 ASVPGTTTDDGMDPGVV--ATTSVTAENSSASIAATAGIGGPPQVDDQETW-RINFYY-- 1839

Db 1957 LSVYNNL-INPFGSTSAIQVTVETRPSEDFEFVIMIRAPSSKTVDSISPAGLTTTPVLT- 2014  
Qy 235 EEMNSRFPPIPLEKUYTPSSAFVQPNQNG---RCTTGGVLGTTQ-LSAVNICTFRGDV 290  
Db 2015 -----GVGNDNRWG--QIVGLQPVGGFSTCNRHWNLANGSTVGWSSPRF-----ADI 2060  
Qy 291 THAGSHDYTMNLAS---QNW-SNYDPTBEIP-----APLGTDP----- 325  
Db 2061 DHRGASYPGSGNATNVLFQWYANAGSAVDNPISQVAPDGPFDMSFVPPNGPGIPAAAGV 2120  
Qy 326 -----FVGKIQGMLTOTTRDGDSTRAHKA--TVSTGS----- 355  
Db 2121 GFCAIWNNSGAPNVTTVQAYELGATGAPGNLQPTTNSGAQTVAKSIYAVVTGAQNP 2180  
Qy 356 -----VHFTPKLGSVQYTTDTNNDFQT-----GQNTKFTPVGVIIQ----- 390  
Db 2181 AGLFVWASGVISTPNAIAITYTPQPDRIVTTPTGTPAAAPVGNKTPIMPASVVRRTGDVNA 2240  
Qy 391 ----DGNHQNQEPQWVLPNYSRGTGHNHVLAPAVAPTTPGEQLLFFRSTMPCCSGYPNNV 447  
Db 2241 TAGSANGTQYGTGSOPLPVTIIGLSLNNYSSA-----LMPGQ---FFVMQLTPASGFMEIG 2292  
Qy 448 L 448  
Db 2293 L 2293  
RESULT 9  
VCWVF9  
coat protein - feline calicivirus (strain F9)  
N/Alternate names: capsid protein  
C:Species: feline calicivirus  
C>Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text\_change 28-Jul-2000  
C:Accession: B43382; P00407; S23702  
R:Cartier, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.  
Virology 190, 443-448, 1992  
A:Title: Identification and sequence determination of the capsid protein gene of feline  
A:Reference number: A45538; MUID:92117861; PMID:1731695  
A:Accession: C45538  
A:Molecule type: genomic RNA; protein  
A:Residues: 1-671 <CAR2>  
A:Cross-references: GB:M86379; NID:G323877; PIDN:AAA79327.1; PID:G323879  
R:Cartier, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.  
Arch. Virol. 122, 223-235, 1992  
A:Title: Identification and sequence determination of the capsid protein gene of feline  
A:Reference number: A45538; MUID:92117861; PMID:1731695  
A:Accession: C45538  
A:Molecule type: genomic RNA; protein  
A:Residues: 1-671 <CAR2>  
A:Cross-references: GB:M86379; NID:G323877; PIDN:AAA79327.1; PID:G323879  
A:Experimental source: strain F9  
A>Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)  
R:Quiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.  
J. Gen. Virol. 73, 2429-2433, 1992  
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel  
A:Reference number: P00407; MUID:93019069; PMID:1402818  
A:Accession: P00407  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>  
A:Cross-references: PIDN:AA823553.1; PID:G257083  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 10.4%; Score 300.5; DB 1; Length 671;  
Best Local Similarity 25.3%; Pred. No. 2.7e-14;  
Matches 141; Conservative 80; Mismatches 238; Indels 99; Gaps 28;  
Qy 12 DGSANLVPEVNEV--MALEPVVGNATAAPVAGQNVDPWIRNFFVQAPGGEFTV--- 66  
Db 127 DGSIT--APEQGTWGGVIAEPASQMSAAOMATKGSVDSEW-----EAPFSFHTSVNW 178

Qy 67 SPRNAPGILMSAPLGPDLPNLYSHLARMYAGGFEVQVILAGNARTAGKIIIFAAVPP 126  
Db 179 STSTQGGKILFKQSLGPLNLYLEHLAKLYVWAGSGIEVRFSGISGSGVFGGKAAIVVPP 238  
Qy 127 NFPTGLESPSOVT--MEPHIIVDVROLEPVLIPDPVNNFNHYNQNSNDSTIKLIAMLY 183  
Db 239 -----GVPVQSTSMQYPHVLFDAQVEPIVFCPLPDLRSLYHLMSDSTDTT-SLIVIMY 292  
Qy 184 TPLRANAGDDVFTVSC--RVLTRPSDDFIFLVPPTVEGRTKFTVPILTVFEEMSNSR 241  
Db 293 NDL-INPYANDANSGCCIVTVETKPGDPFKHLLKPP--GSMLTGHSIPSDLIPTKTS-- 347  
Qy 242 FPIPLEKLYTPSSAFVQNP---QNGR-----CTTGGVLLGTTQLSANNICTFRGDVTHI 293  
Db 348 --LWIGNRYSDITDFVIRPFVQANRHFDENQETAGWSTPRFRISVTITEQNGAKLGI 405  
Qy 294 AGSHDYTMNLASQWNSYD-PTEEIPAPLGPDPFVGKIQGMLTOTTRDGDSTRA--HKAT 350  
Db 406 GVATDIYVGPDPDGPDDTTIFGELIPA-----GDYA-----ITNGTGNDITATGYDTAD 455  
Qy 351 VSTGSVHF-----TPKLGSVQYTT-----DTNNDFQTGQNTKFTPVGVIIQ 390  
Db 456 IIKNTNFRGMVYCGSLQORAGDKKISNTAFITTATLOGDNNKINPCNTIDQSKIIVVQ 515  
Qy 391 D---GNHQNQEPQWVLPNYSG-----RTGHNHVLAPAVAPTTP-----GEQLLFFRSTM 437  
Db 516 DNHVGKKAQTSDDTLALLGYTGIGQAIGSDRDRVVRIS-TLPETGARGGNHPIFYKNSI 574  
Qy 438 PGCSYPMNMLDCLLPQEMVQHFQCEAA-----PAQSDVALLRFVNPDTGRVLFECKLH 491  
Db 575 K--LGIVRSIDVFNQ--ILHTRQSLNHYLLPPDS--FAYRII---DSNGSWFDIGID 627  
Qy 492 KSGYVTVTAHTGPHDIVIP 509  
Db 628 SDGFSFVGVSGFGKLEPP 645  
RESULT 10  
A48562  
coat protein - San Miguel sea lion virus (serotype 1)  
N/Alternate names: capsid protein  
C:Species: San Miguel sea lion virus  
C>Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
C:Accession: A48562  
R:Neill, J.D.  
Virus Res. 24, 211-222, 1992  
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel  
eins.  
A:Reference number: A48562; MUID:92410750; PMID:1529644  
A:Accession: A48562  
A:Molecule type: genomic RNA  
A:Residues: 1-702 <NET>  
A:Cross-references: GB:M87481; NID:G334882; PIDN:AAA16217.1; PID:G334884  
A>Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBIP:113565)  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 10.3%; Score 296.5; DB 1; Length 702;  
Best Local Similarity 26.1%; Pred. No. 5.7e-14;  
Matches 141; Conservative 72; Mismatches 199; Indels 129; Gaps 30;  
Qy 11 SDG-STANLVPEVNEVMALEPVVVG-----AAIAAPVAGQNVDPWIRNFFVQAPGGEFT 65  
Db 153 SDGPGGADIVTEEGTQVTVQOQPVPAQASALTTLAAASTGK--TVD-----CEWT 198  
Qy 66 V-----SPRNAPGILMSAPLGPDLPNLYSHLARMYAGGFEVQVILAGNART 115  
Db 199 TFFSYHTAVNNTTEAQCKILFSRALSPELNPYLRHISLSYSTWSGGIDVRFVTSGSGVF 258  
Qy 116 AKGIIFAAPVNP-PTEGLSPQVMTMFPHIIIVDROLEPVLIPDPVNNFNHYNQNSD 174  
Db 259 GKKLAALIVPPGIEBEV--SPTML-QYPHVLFDARQTEPFVITPIDIRKTLYHSMDDTDT 315

```
QY 175 TIKLAMYTLPL-RANNAGDDVFTVSRVLRPSDDFPIFLVPTVESRTPKPTVPILT 233
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
316 T-LRVIMVYNELINPEYEQSEPKSCSITVETRPSSDFTFSLKKPP--GSLLKHSIPSDL 372
QY 234 VEEMNSRFPPILEKLYICPSSAFVVOQO---NGRCITDGVLLGTQLSVAVNICTPRGD 289
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
373 IPR--NSRH--WMGNRWSTIDGFVVQPRVFNOSNRHFDFTSTTGWSTPYIPI----- 422
QY 290 VTHIAGSHDYTNLASQWNSNDPTTEE---IPA-PLGTDPFVGKIQKMLTQTTR-EDGSTRAHKAT 343
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
423 -----EVILEKDRGGQYFKVTDTEKSLVPLGPDGWPDTT-----IPTAMTASNGNY 469
QY 344 --TRAHATVSTGSHVH-----TPKLSGVQVYTTDNDPFGQNTKFTFPVGIODG 392
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
470 DYTVAEYRITNNGT-HFKGYFMGNLTTKVG-----SDNLGETQQTSTLFAVSNGYKQD 524
QY 393 N-----NHQNEPQOWLPNYSGRGTGHNHVLAPAVPTFPGEQLLFFRSTMPGCGYPNNML 448
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
525 NTINPTHKITSLSLVYD-----ANNVSAATAKTIT-----WHSTM-----SHL 563
QY 449 DCLLPQEWYQHFCQEAAPQSD-----VALL--RFVNPDTGRVLPECKLHKSQVTVVAH 500
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
564 GYVLVDE-----SPVGSDBTKVRIATLPEAFTNGGNFPVFTNKI-QIGHFDRAH 613
QY 501 T 501
Db 614 T 614
```

## RESULT 11

VCWVFF

coat protein - feline calicivirus (strain Japanese F4)

N;Alternate names: capsid protein

C;Species: feline calicivirus

C;Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000

C;Accession: B40481

R;Tohya, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamaza

Virology 183, 810-814, 1991

A;Title: Sequence analysis of the 3'-end of feline calicivirus genome.

A;Reference number: A40481; MUID:91306470; PMID:1853578

A;Molecule type: genomic RNA

A;Residues: 1-668 &lt;TOH&gt;

A;Cross-references: GB:D90357; NID:G221264; PID:BAA14371.1; PID:G221266

C;Superfamily: feline calicivirus coat protein

C;Keywords: capsid protein; coat protein; Glycoprotein

F;177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.2%; Score 294; DB 1; Length 668;

Best Local Similarity 24.6%; Pred. No. 8.1e-14;

Matches 120; Conservative 61; Mismatches 171; Indels 136; Gaps 21;

QY 12 DGSTANLVPEYNNEV--MALEPVVGAIAAPVAGQQNVDPWIRNNFVQAPGGEFTV--- 66

Db 127 DGSIT--TPEGTVAGGVIAEPSAQMSTADMASGKSDSEW-----EAFPSFTSVNW 178

QY 67 SPRNAPGEILWSAPLGDPLNLYLSHLMRYNGYAGGFEVQVILAGNAFTAGKIIFAAVPP 126

Db 179 STSETQKILFKQSLGLPLNLYLSHLSKLYVWAGSIEVRFSGVFGKLAIVVPP 238

QY 127 NFPTGELSPSQVTV--MPPHIVDVRLQLEPLIPDPVRNNFHYNQSNDSSTIKLIAMLY 183

Db 239 -----GVDPVQSTSMLOYPHVLFDARQVEPIFIPDLRSLYHVMSDITDTT-SLIVMY 292

QY 184 TPLRANNAGDDVFTVSC--RVLTRPSDDFPIFLVPP-----TVES----- 222

Db 293 NDL-INPYANDSNSGGCIWTVETKPGDPKFKHLKKPPGSLVTHGSIPSDLIPIKSSSIWIG 351

QY 223 -----RTKPFVTPILTVEEMNSRFPPIPLE 247

Db 352 NRYWTDITDFVIRPVFVFOANRHFDPNQETAGWSIFRFPRII---TISEKNGSKLGIGVA 408

```
QY 248 KLYTGP-----SSAFVVPQNGRCRTDGVLLGTTQ---LSAV--NICTFRGDVT 291
Db 409 TDYIIIGIPDGMVDTTIIADKLIPAGDYSITTCGENDIKTAQAYDTAAVVKNTNFRG--M 466
QY 292 HIAGSHDYTNLASQWNSNDPTTEEIPAPLGTDPFVGKIQKMLTQTTR-EDGSTRAHKAT 350
Db 467 YICGS-----LQRAWGD-----KKISNTAFITTAIRDGNEIKPSNT 502
QY 351 VSTGSHVHTPKLSGVQVYTTDNDNFOTQWNT-----KFTPVGVIQDGNHQNPEQWVLPN 406
Db 503 IDMTKL-----AVQDTHVEQVQTSDDTLALLGYTGIGBEAIGSNRDRVVRISVLPE 555
QY 407 YSGRTGHN 414
Db 556 AGARGGNH 563

RESULT 12
JQ2354
capsid protein - feline calicivirus (strain NADC)
C;Species: feline calicivirus
C;Date: 28-Aug-1995 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C;Accession: JQ2354
J;Seal, B.S.; Kidpath, J.F.; Mengeling, W.L.
R;Gen. Virol. 74, 2519-2524, 1993
A;Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A;Reference number: JQ2354; MUID:94065683; PMID:7504075
A;Accession: JQ2354
A;Molecule type: mRNA
A;Residues: 1-668 <SEA>
A;Cross-references: GB:I09718; NID:G305104; PID:AAA16485.1; PID:G305105
C;Superfamily: feline calicivirus coat protein
C;Keywords: capsid protein; coat protein

Query Match 10.2%; Score 293.5; DB 2; Length 668;
Best Local Similarity 26.3%; Pred. No. 8.9e-14;
Matches 109; Conservative 56; Mismatches 196; Indels 53; Gaps 16;

QY 32 VVGAIAAPVAGQQNVDPWIRNNFVQAPGGEF-----TVSPRNAPGEILWSAPLGD 84
Db 138 VVGVIAEPSAQMSTAAD-MATGKSDSEWFAFFSFTSVNMWSTSETQKILFKQSLGPL 196
QY 85 LNPYLHLMRYNGYAGGFEVQVILAGNAFTAGKIIFAAVPPNFPTEGLSPQVT---MF 141
Db 197 LNPYLEHLSKLYVWAGSIEVRFSGVFGKLAIVVPP-----GVDPVQSTSMLOQ 251
QY 142 PHIIIVDVRLQLEPLIPDPVRNNFHYNQSNDSSTIKLIAMLYTFLRANNAGDDVFTVSC- 200
Db 252 PHVLFDARQVDFVIFSIDLRLSTLYHLMPTDIT-SLIVIMVNDL-INPYANDSNSGCI 309
QY 201 -RVLTRPSDDFPIFLVPTVESRTPKPTVILTVEEMNSRFPPIPLEKLYTGPSAFVW 259
Db 310 VTVETKPGDPKFKHLKKPP--GSMITHGSPVSDLIPIKSSS-----LWIGNRWSDITDFW 363
QY 260 QP-----QNGRCITDGVLLGTQLSVAVNICTFRGDVTHIAGSHDYTNLASQWNSNY 311
Db 364 RPFVFOANRHFNFQETAGWSAPRFRPITIISEKSKGLGIGVATDVIIVGIPDGWPD 423
QY 312 DTEETIPAPLGPDPF-VGKIQKMLTQTTRPDGSTRAHKATVSTGSHVH-----TPK 361
Db 424 TTAEDL-TPAG--DYAITSGNGNDITGTSEYDSTEVIKNNTNFRGMVTCGSLQRAWGDKK 480
QY 362 LGSVQVTTDT---NNDFTQGTQKFTFPVGIQD---GNHQNPEQWVLPNYSG 409
Db 481 ISNTAFITTAIEGKNKIRPSNTIDMTKIAVQDTHVGEVQTSDDALLGTYG 534

RESULT 13
VCWVFF
coat protein - feline calicivirus (strain CFI/68 FIV)
N;Alternate names: capsid protein
C;Species: feline calicivirus
C;Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 20-Aug-1999
```



[illegible]



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 30.8574 Seconds  
(without alignments)  
5419.275 Million cell updates/sec

Title: US-09-926-799-2  
Perfect score: 2819  
Sequence: 1 MMASKDTSVVDGASGAGQ.....YOLKPVGTASSARGRLGLRR 530

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : SPTREMBL\_25.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phage.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriaph.\*
  - 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2819	100.0	530	12 Q9QT39	Q9QT39 norwalk-lik
2	2789	98.9	530	12 Q83884	Q83884 norwalk vir
3	2737	97.1	517	12 Q9JH41	Q9JH41 norwalk-lik
4	2713	96.2	530	12 Q83876	Q83876 norwalk vir
5	2093	74.2	544	12 Q9DU46	Q9DU46 chiba virus
6	2091	74.2	544	12 Q9QTE7	Q9QTE7 chiba virus
7	2084	73.9	544	12 Q91I85	Q91I85 human calic
8	2081	73.8	544	12 Q918C5	Q918C5 norwalk-lik
9	2076	73.6	544	12 Q9IEU2	Q9IEU2 human calic
10	2063	73.2	544	12 Q918D1	Q918D1 norwalk-lik
11	2046	72.6	544	12 Q91V41	Q91V41 human calic
12	2044.5	72.5	540	12 Q9YS14	Q9YS14 norwalk vir
13	2044	72.5	544	12 Q91V36	Q91V36 human calic
14	2031	72.0	531	12 Q8V7J5	Q8V7J5 norwalk-lik
15	2029.5	72.0	546	12 Q8JW44	Q8JW44 norwalk-lik
16	2027.5	71.9	540	12 Q8BCA5	Q8BCA5 human calic

17	2023.5	71.8	543	12 Q8JXJ0	Q8JXJ0 norwalk-lik
18	2019.5	71.6	546	12 Q91V47	Q91V47 human calic
19	2017.5	71.6	544	12 Q8VA27	Q8VA27 human calic
20	1997.5	70.9	543	12 Q91V43	Q91V43 human calic
21	1985.5	70.4	543	12 Q918B9	Q918B9 norwalk-lik
22	1964	69.7	539	12 Q8JVU5	Q8JVU5 norwalk-lik
23	1950.5	69.2	544	12 Q66418	Q66418 desert shie
24	1949	69.1	545	12 Q8VA02	Q8VA02 human calic
25	1939	68.8	541	12 Q91V42	Q91V42 human calic
26	1930	68.5	545	12 Q918C8	Q918C8 norwalk-lik
27	1930	68.5	545	12 Q9WI82	Q9WI82 norwalk-lik
28	1930	68.5	545	12 Q918C2	Q918C2 norwalk-lik
29	1914	67.9	539	12 Q91V48	Q91V48 human calic
30	1904	67.5	545	12 Q91H12	Q91H12 human calic
31	1902	67.5	545	12 Q91V45	Q91V45 human calic
32	1832	65.0	543	12 Q8BCA3	Q8BCA3 human calic
33	1279.5	45.4	522	12 Q8B558	Q8B558 bovine ente
34	1275.5	45.2	522	12 Q8B4Z5	Q8B4Z5 norwalk-lik
35	1273.5	45.2	522	12 Q8B4Y9	Q8B4Y9 norwalk-lik
36	1271.5	45.1	520	12 Q7TBL1	Q7TBL1 bovine ente
37	1268.5	45.0	520	12 Q7TBL7	Q7TBL7 bovine ente
38	1266.5	44.9	522	12 Q8V629	Q8V629 norwalk-lik
39	1260.5	44.7	522	12 Q8V628	Q8V628 norwalk-lik
40	1257.5	44.6	522	12 Q8B4Z2	Q8B4Z2 norwalk-lik
41	1254	44.5	519	12 Q9YQ22	Q9YQ22 bovine cali
42	1243	44.1	542	12 Q913B6	Q913B6 human calic
43	1238	43.9	542	12 Q66293	Q66293 calicivirus
44	1235	43.8	540	12 Q917X1	Q917X1 norwalk-lik
45	1233	43.7	548	12 Q915C7	Q915C7 human calic

ALIGNMENTS

RESULT 1

Q9QT39	PRELIMINARY;	PRT;	530 AA.
AC	Q9QT39;		
DC	01-MAY-2000 (TREMREL. 13, Created)		
DT	01-MAY-2000 (TREMREL. 13, Last sequence update)		
DT	01-OCT-2003 (TREMREL. 25, Last annotation update)		
DE	Capsid protein.		
OS	Norwalk-like virus genogroup 2.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_TaxID=122929;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Aichi124-89;		
RX	MEDLINE=20440685; PubMed=10970413;		
RA	Kobayashi S., Sakae K., Suzuki Y., Shinozaki K., Okada M., Ishiko H.,		
RA	Kanada K., Suzuki K., Natori K., Miyamura T., Takeda N.;		
RT	"Molecular Cloning, Expression, and Antigenicity of Seto Virus		
RT	Belonging to Genogroup 1 Norwalk-Like Viruses.";		
RL	31-03-2004-1603-3383492-33994(2000)15		
DR	EMBL; AB031013; BAA83413.1; -		
DR	InterPro; IPR004005; Calici_coat.		
DR	InterPro; IPR008975; Viral_cap_coat.		
DR	Pfam; PF00915; Calici_coat_1_coat.		
SQ	SEQUENCE 530 AA; 56443 MW; 24845684A6902300 CRC64;		
Query Match	100.0%; Score 2819; DB 12; Length 530;		
Best Local Similarity	100.0%; Pred.No.1.5e-221;		
Matches 530;	Conservative 0; Mismatches 0; Indels 0; Gaps 0;		

QY	1	MMASKDTSVVDGASGAGQLVPEVNASDPVAGSSTAVATAGQVNPIDPFIINN	60
Db	1	MMASKDTSVVDGASGAGQLVPEVNASDPVAGSSTAVATAGQVNPIDPFIINN	60
QY	61	VOAPQGETTISPNNTPGGVLFDSLSGPHLNPFLHLHLSOMYNGWGNMVRIMLAGNAFTA	120
Db	61	VOAPQGETTISPNNTPGGVLFDSLSGPHLNPFLHLHLSOMYNGWGNMVRIMLAGNAFTA	120

```
QY 121 GKIIIVSCIPPGSGSHNLTTAQTALFPHVIADVRTLDPIEVLPLEDVNRNVLFHNNDRNQTM 180
DB 121 GKIIIVSCIPPGSGSHNLTTAQTALFPHVIADVRTLDPIEVLPLEDVNRNVLFHNNDRNQTM 180
QY 181 RLVCMLYTLPLRTGGTGDSEFVVGAVRMTCPSPDENFLFVLPPTVEQKTRPFTLNLPLSS 240
DB 181 RLVCMLYTLPLRTGGTGDSEFVVGAVRMTCPSPDENFLFVLPPTVEQKTRPFTLNLPLSS 240
QY 241 LNSRAPLPIISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
DB 241 LNSRAPLPIISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
QY 301 LTELDTGTPPHPEGPAPIGFDPDLGGCDWHNNMTQFCHSSQTYDVTDPDTFVPHLGSIQ 360
DB 301 LTELDTGTPPHPEGPAPIGFDPDLGGCDWHNNMTQFCHSSQTYDVTDPDTFVPHLGSIQ 360
QY 361 ANIGSGNIVGLSWSPSPSHSGSOVDLWKIPNYGSSITEATHLAPSVPYPGFGEVLVF 420
DB 361 ANIGSGNIVGLSWSPSPSHSGSOVDLWKIPNYGSSITEATHLAPSVPYPGFGEVLVF 420
QY 421 FMSKIPGPGAYSILPCLLPQEIYISHLASEQAPTVEAALLHYVDPTGRTLGEFKAYPDGF 480
DB 421 FMSKIPGPGAYSILPCLLPQEIYISHLASEQAPTVEAALLHYVDPTGRTLGEFKAYPDGF 480
QY 481 LTCVPNGASSGPOOLPINGVVFVSWSRFYQLKPVGTASSARGRLGLRR 530
DB 481 LTCVPNGASSGPOOLPINGVVFVSWSRFYQLKPVGTASSARGRLGLRR 530
QY 481 LTCVPNGASSGPOOLPINGVVFVSWSRFYQLKPVGTASSARGRLGLRR 530
DB 481 LTCVPNGASSGPOOLPINGVVFVSWSRFYQLKPVGTASSARGRLGLRR 530
RESULT 2
Q83884 PRELIMINARY; PRT; 530 AA.
AC Q83884;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE 58 kDa capsid protein.
OS Norwalk virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=9303935; PubMed=8391187;
RA Jiang X., Wang M., Wang K., Estes M.K.;
RT "Sequence and genomic organization of Norwalk virus,";
RL Virology 195:51-61(1993).
DR EMBL: M87661; AAB50466.1; -.
DR PIR: B37471; B37471.
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
SQ SEQUENCE 530 AA; 56571 MW; CRED8D86366CD236 CRC64;
Query Match 98.9%; Score 2789; DB 12; Length 530;
Best Local Similarity 98.5%; Pred. No. 4.le-219;
Matches 522; Conservative 5; Mismatches 3; Indels 0; Gaps 0;
QY 1 MMASKDATSVDGASGAGQLVPEVNASDPLAMPDVAGSSTAVATAGVNPIDPWIINN 60
DB 1 MMASKDATSVDGASGAGQLVPEVNASDPLAMPDVAGSSTAVATAGVNPIDPWIINN 60
QY 61 VQAPQGETIIPNNTPGCVLFDLSLGLPHNLPFLHLSQWYNGWGNMVRIMLAGNAFTA 120
DB 61 VQAPQGETIIPNNTPGCVLFDLSLGLPHNLPFLHLSQWYNGWGNMVRIMLAGNAFTA 120
QY 121 GKIIIVSCIPPGSGSHNLTTAQTALFPHVIADVRTLDPIEVLPLEDVNRNVLFHNNDRNQTM 180
DB 121 GKIIIVSCIPPGSGSHNLTTAQTALFPHVIADVRTLDPIEVLPLEDVNRNVLFHNNDRNQTM 180
QY 181 RLVCMLYTLPLRTGGTGDSEFVVGAVRMTCPSPDENFLFVLPPTVEQKTRPFTLNLPLSS 240
DB 181 RLVCMLYTLPLRTGGTGDSEFVVGAVRMTCPSPDENFLFVLPPTVEQKTRPFTLNLPLSS 240
```

```
QY 241 LNSRAPLPIISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
DB 241 LNSRAPLPIISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
QY 301 LTELDTGTPPHPEGPAPIGFDPDLGGCDWHNNMTQFCHSSQTYDVTDPDTFVPHLGSIQ 360
DB 301 LTELDTGTPPHPEGPAPIGFDPDLGGCDWHNNMTQFCHSSQTYDVTDPDTFVPHLGSIQ 360
QY 361 ANIGSGNIVGLSWSPSPSHSGSOVDLWKIPNYGSSITEATHLAPSVPYPGFGEVLVF 420
DB 361 ANIGSGNIVGLSWSPSPSHSGSOVDLWKIPNYGSSITEATHLAPSVPYPGFGEVLVF 420
QY 421 FMSKIPGPGAYSILPCLLPQEIYISHLASEQAPTVEAALLHYVDPTGRTLGEFKAYPDGF 480
DB 421 FMSKIPGPGAYSILPCLLPQEIYISHLASEQAPTVEAALLHYVDPTGRTLGEFKAYPDGF 480
QY 481 LTCVPNGASSGPOOLPINGVVFVSWSRFYQLKPVGTASSARGRLGLRR 530
DB 481 LTCVPNGASSGPOOLPINGVVFVSWSRFYQLKPVGTASSARGRLGLRR 530
RESULT 3
Q9JH41 PRELIMINARY; PRT; 517 AA.
AC Q9JH41;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein (fragment).
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20150145; PubMed=10686033;
RA Yoda T., Terano Y., Shimada A., Suzuki Y., Yamazaki K., Sakon N.,
RA Oishi I., Utagawa E.T., Okuno Y., Shibata T.;
RT "Expression of recombinant Norwalk-like virus capsid proteins using a
bacterial system and the development of its immunologic detection.";
RL J. Med. Virol. 60:475-481(2000).
DR EMBL: AB028247; BAA89031.2; -.
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
FT NON_TER 1 517
FT NON_TER 517 517
SQ SEQUENCE 517 AA; 55011 MW; BAD7DF35E1DD0DF6 CRC64;
Query Match 97.1%; Score 2737; DB 12; Length 517;
Best Local Similarity 99.2%; Pred. No. 7e-215;
Matches 513; Conservative 1; Mismatches 3; Indels 0; Gaps 0;
QY 8 ATSSVDGASGAGQLVPEVNASDPLAMPDVAGSSTAVATAGVNPIDPWIINN 67
DB 1 ATSSVDGASGAGQLVPEVNATSDPLAMPDVAGSSTAVATAGVNPIDPWIINN 60
QY 68 FTISPNNTPGGVLFDLISLGLPHNLPFLHLSQWYNGWGNMVRIMLAGNAFTAGKIIVSC 127
DB 61 FTISPNNTPGGVLFDLISLGLPHNLPFLHLSQWYNGWGNMVRIMLAGNAFTAGKIIVSC 120
QY 128 IPPFGSHNLTTAQTALFPHVIADVRTLDPIEVLPLEDVNRNVLFHNNDRNQTMRLVCMY 187
DB 121 IPPFGSHNLTTAQTALFPHVIADVRTLDPIEVLPLEDVNRNVLFHNNDRNQTMRLVCMY 180
QY 188 TPLRTGGTGDSEFVVGAVRMTCPSPDENFLFVLPPTVEQKTRPFTLNLPLSSLSNSRAP 247
DB 181 TPLRTGGTGDSEFVVGAVRMTCPSPDENFLFVLPPTVEQKTRPFTLNLPLSSLSNSRAP 240
QY 248 LPISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVINLDELDT 307
```

```
Db 241 LPISMGISPNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGTSNGTVINLTDLGT 300
Qy 308 PFHPEGPAPICFPDGLGCDWHINNTQFHSSTQYDVTTPDTFVPHLGSIQANGISG 367
Db 301 PFHPEGPAPICFPDGLGCDWHINNTQFHSSTQYDVTTPDTFVPHLGSIQANGISG 360
Qy 368 NYIGVLSWVSPSPSPSHPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGVLPVPMKIPG 427
Db 361 NYIGVLSWVSPSPSPSHPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGVLPVPMKIPG 420
Qy 428 PGAYSLPCLLPQYIYSHLASQAPTVGEAALLHYVDPTGRTGLGFKAYPDGFLTCVPNG 487
Db 421 PGAYSLPCLLPQYIYSHLASQAPTVGEAALLHYVDPTGRTGLGFKAYPDGFLTCVPNG 480
Qy 488 ASSGPGQQLPINGVFVFSWVSRYQLKPVGTASSARG 524
Db 481 ASSGPGQQLPINGVFVFSWVSRYQLKPVGTASTARG 517

RESULT 4
ID Q83876 PRELIMINARY; PRT; 530 AA.
AC Q83876;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Norwalk virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SRV-KY-89/89/J;
EX MEDLINE=94335115; PubMed=8057474;
RA Wang J., Jiang X., Madore H.P., Gray J., Desselberger U., Ando T.,
RA Sato Y., Oishi I., Lew J.F., Green K.Y., et al.;
RT "Sequence characterization of small, round-structured viruses in the Norwalk
RL virus group." Virology 68:5980-5990(1994).
DR EMBL; L23828; AA59229.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 530 AA; 56402 MW; 9C0E2C3199CCF051 CRC64;

Query Match 96.2%; Score 2713; DB 12; Length 530;
Best Local Similarity 96.8%; Pred. No. 6.6e-213;
Matches 513; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

Qy 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIINN 60
Db 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIINN 60
Qy 61 VOAPQGEFTISPNPTPGDVLFDLSGLPHLNPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
Db 61 VOAPQGEFTISPNPTPGDVLFDLSGLPHLNPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
Qy 121 GKIIVSCIPPGGSHNLITIAQATLPPHVIADVRTLDPIEVLDPVFNHNDNQTM 180
Db 121 GKIIVSCIPPGGSHNLITIAQATLPPHVIADVRTLDPIEVLDPVFNHNDNQTM 180
Qy 181 RLVCMLYPLRTGGTGSFVVGAVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPLSS 240
Db 181 RLVCMLYPLRTGGTGSFVVGAVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPLSS 240
Qy 241 LNSNRAPLPISGMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
Db 241 LNSNRAPLPISGMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGTSNGTVIN 300
Qy 301 LTDLGTPFHPFEGPAPICFPDGLGCDWHINNTQFHSSTQYDVTTPDTFVPHLGSIQ 360
Db 301 LTDLGTPFHPFEGPAPICFPDGLGCDWHINNTQFHSSTQYDVTTPDTFVPHLGSIQ 360
```

```
Qy 361 ANGISGNYIGVLSWVSPSPSHPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGVLPV 420
Db 361 ANGISGNYIGVLSWVSPSPSHPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGVLPV 420
Qy 421 FMSKIPGPGAYSLPCLLPQYIYSHLASQAPTVGEAALLHYVDPTGRTGLGFKAYPDG 480
Db 421 FMSKIPGPGAYSLPCLLPQYIYSHLASQAPTVGEAALLHYVDPTGRTGLGFKAYPDG 480
Qy 481 LTCVPNGASSGPGQQLPINGVFVFSWVSRYQLKPVGTASSARGRLGLRR 530
Db 481 LTCVPNGASSGPGQQLPINGVFVFSWVSRYQLKPVGTASTARGRLGLRR 530

RESULT 5
ID Q9DU46 PRELIMINARY; PRT; 544 AA.
AC Q9DU46;
DT 01-MAR-2001 (TremBLrel. 16, Created)
DT 01-MAR-2001 (TremBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Chiba 407/1987/JP;
RA Someya Y., Takeda N., Miyamura T.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Chiba 407/1987/JP;
EX MEDLINE=20569531; PubMed=11118371;
RA Someya Y., Takeda N., Miyamura T.;
RT "Complete nucleotide sequence of the Chiba virus genome and functional
RT expression of the 3C-like protease in Escherichia coli."
RL Virology 278:490-500(2000).
DR EMBL; AB042808; BAB18267.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58351 MW; 1DB7A4E81FBB246F CRC64;

Query Match 74.2%; Score 2093; DB 12; Length 544;
Best Local Similarity 72.7%; Pred. No. 3e-162;
Matches 396; Conservative 53; Mismatches 80; Indels 16; Gaps 8;

Qy 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIINN 60
Db 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIINN 60
Qy 61 VOAPQGEFTISPNPTPGDVLFDLSGLPHLNPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
Db 61 VOAPQGEFTISPNPTPGDVLFDLSGLPHLNPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
Qy 121 GKIIVSCIPPGGSHNLITIAQATLPPHVIADVRTLDPIEVLDPVFNHNDNQTM 180
Db 121 GKIIVSCIPPGGSHNLITIAQATLPPHVIADVRTLDPIEVLDPVFNHNDNQTM 180
Qy 181 RLVCMLYPLRTGGTGSFVVGAVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPL 238
Db 181 RLVCMLYPLRTGGTGSFVVGAVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPL 238
Qy 239 SSLNSRAPLPISGMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRG--TSNG 296
Db 240 KYLSNSRIPNPIEGMSLSPDQTNVQFQNGRCTLGRLVGTTPVSVLSOLCKFRGRITSQ 299
Qy 297 TVINLTDLGTPFHPFEGPAPICFPDGLGCDWHINNTQFHSSTQYDVTTPDTFVPHLGSIQ 351
Db 300 RVLNLTDLGTPFHPFEGPAPICFPDGLGCDWHINNTQFHSSTQYDVTTPDTFVPHLGSIQ 359
```

```
QY 352 FVPHLGSII--QANGIGSGNYIGVLVSWVSPSHPSGSQVDLWKIPNYGSSITEATHLAPSV 409
Db 360 FVPHLSSITLDENVSSGGDIYIGTQWTSPPSDSGGANTFNFKIPDYGSSLAESAQLAPAV 419
QY 410 YPGFGEVLVFMSPKIPGP--GAYSL-PCLLPQEIYISHLASEQAQPTVGEAALLHYVDPD 465
Db 420 YPGFNEVIVFMASIPGNQSGSNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPD 479
QY 466 TGRTLGEFKAYPDGFLTCVPGNGASSGPOQLPINGVVFVSWVSRFYQLKPVGTASSARGR 525
Db 480 TNRNLGEFKLYPGGYLTCVPSNSSSTGPQQLPDGVFVFASWVSRYQLKPVGTAGPARGR 539
QY 526 LGRR 530
Db 540 LGRR 544

RESULT 6
Q9QTE7 PRELIMINARY; PRT; 544 AA.
AC Q9QTE7;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Chiba 407;
RA Natori K., Suzuki K., Yamakawa Y., Tateumi M., Sakae K., Kobayashi S.,
RA Shinozaki K., Ishiko H., Miyamura T., Takeda N.;
RT "Expression and self-assembly of capsid proteins of the Chiba virus, a
RT genetically distinct Norwalk-like virus.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB028679; BAA82106.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58323 MW; ABDIC1FC4F93D872 CRC64;

Query Match 74.2%; Score 2091; DB 12; Length 544;
Best Local Similarity 72.74; Pred. No. 4.3e-162;
Matches 396; Conservative 53; Mismatches 80; Indels 16; Gaps 8;

QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMPDVAGSSTAVATAGQVNPIDPWIIINF 60
Db 1 MMASKDATPSADGATGAGQLVPEVNTADIPIDPVAGSSTALATAGQVNLIDPWIIINF 60
QY 61 VQAPQGEFTTSPNNTPGGVLFDSLGLPHLNPFLHLSQMYNGVGNMVRVIMLAGNAFTA 120
Db 61 VQAPQGEFTTSPNNTPGDVLFDLQGLPHLNPFLHLSQMYNGVGNMVRVIMLAGNAFTA 120
QY 121 GKIIIVCIPPGFSGSHNLITIAQATLFPHVIAVRTLDPIEVLDPVLEDRVNLPHNNDRNQTM 180
Db 121 GKIIICVPPGQFQSRITLSIAQATLFPHVIAVRTLDPIEVLDPVLEDRVNLPHNNDRNQTM 179
QY 181 RLVCMLYTPRLTGGGTG--DSFVAVGRVMTCPSPDFNLFVLPVPTVEQKTRPFTLNLPL 238
Db 181 RLVCMLYTPRLTGGGTG--DSFVAVGRVMTCPSPDFNLFVLPVPTVEQKTRPFTLNLPL 239
QY 239 SSLNSRAPLPIISGMISPDNVQSVQFONGRCTLDGLRGVLTTPVLSLHVAKIRG--TSNG 296
Db 240 KYLSNRIPNPIEGMSLSPDQTNVQFNGRCTIDGQPLGTTPVSVSQLCKFRGRTISGQ 299
QY 297 TVINLTDLGTPHPFEGPAPIGFPPDLGGCDWHINMTQFGHSSQTDVDTT---PDT-- 351
Db 300 RVLNLTDLGSPFGGAPAPAGFPDLGSCDWHIEMSKIPNSQTNNPIVTSVKPNSQ 359
QY 352 FVPHLGSII--QANGIGSGNYIGVLVSWVSPSHPSGSQVDLWKIPNYGSSITEATHLAPSV 409
Db 360 FVPHLSSITLDENVSSGGDIYIGTQWTSPPSDSGGANTFNFKIPDYGSSLAESAQLAPAV 419
QY 410 YPGFGEVLVFMSPKIPGP--GAYSL-PCLLPQEIYISHLASEQAQPTVGEAALLHYVDPD 465
Db 420 YPGFNEVIVFMASIPGNQSGSNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPD 479
QY 466 TGRTLGEFKAYPDGFLTCVPGNGASSGPOQLPINGVVFVSWVSRFYQLKPVGTASSARGR 525
Db 480 TNRNLGEFKLYPGGYLTCVPSNSSSTGPQQLPDGVFVFASWVSRYQLKPVGTAGPARGR 539
QY 526 LGRR 530
Db 540 LGRR 544
```

```
Db 360 FVPHLSSITLDENVSSGGDIYIGTQWTSPPSDSGGANTFNFKIPDYGSSLAESAQLAPAV 419
QY 410 YPGFGEVLVFMSPKIPGP--GAYSL-PCLLPQEIYISHLASEQAQPTVGEAALLHYVDPD 465
Db 420 YPGFNEVIVFMASIPGNQSGSNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPD 479
QY 466 TGRTLGEFKAYPDGFLTCVPGNGASSGPOQLPINGVVFVSWVSRFYQLKPVGTASSARGR 525
Db 480 TNRNLGEFKLYPGGYLTCVPSNSSSTGPQQLPDGVFVFASWVSRYQLKPVGTAGPARGR 539
QY 526 LGRR 530
Db 540 LGRR 544

RESULT 7
Q91I85 PRELIMINARY; PRT; 544 AA.
AC Q91I85;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Koblenz/433/2000/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=165252;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Koblenz 433/2000/DE;
RA Kuenkel U., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RT with Norwalk-like viruses in Germany.";
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF394960; AAK72048.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58222 MW; 1B39A595E5733A6A8 CRC64;

Query Match 73.9%; Score 2084; DB 12; Length 544;
Best Local Similarity 72.5%; Pred. No. 1.6e-161;
Matches 396; Conservative 51; Mismatches 81; Indels 18; Gaps 8;

QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMPDVAGSSTAVATAGQVNPIDPWIIINF 60
Db 1 MMASKDATPSADGATGAGQLVPEVNTAGPIPIDPVAGSSTALATAGQVNLIDPWIIINF 60
QY 61 VQAPQGEFTTSPNNTPGGVLFDSLGLPHLNPFLHLSQMYNGVGNMVRVIMLAGNAFTA 120
Db 61 VQAPQGEFTTSPNNTPGDVLFDLQGLPHLNPFLHLSQMYNGVGNMVRVIMLAGNAFTA 120
QY 121 GKIIIVCIPPGFSGSHNLITIAQATLFPHVIAVRTLDPIEVLDPVLEDRVNLPHNNDRNQTM 180
Db 121 GKIIICVPPGQFQSRITLSIAQATLFPHVIAVRTLDPIEVLDPVLEDRVNLPHNNDRNQTM 179
QY 181 RLVCMLYTPRLTGGGTG--DSFVAVGRVMTCPSPDFNLFVLPVPTVEQKTRPFTLNLPL 238
Db 180 RLLCMLYTPRLTGGASGSDTSFVAVGRVLTCPGPDFNLFVLPVPTVEQKTRPFTVNPIL 239
QY 239 SSLNSRAPLPIISGMISPDNVQSVQFONGRCTLDGLRGVLTTPVLSLHVAKIRG--TSNG 296
Db 240 KYLSNRIPNPIEGMSLSPDQTNVQFNGRCTIDGQPLGTTPVSVSQLCKFRGRTISGQ 299
QY 297 TVINLTDLGTPHPFEGPAPIGFPPDLGGCDWHINMTQFGHSSQTDVDTT---VDTTDP 350
Db 300 KVLNLTDLGSPFMAFAAPAPAGFPDLGSCDWHIEMSKIPNS--TQSNPIVSVSVKPNQ 358
QY 351 TVPHLGSII--QANGIGSGNYIGVLVSWVSPSHPSGSQVDLWKIPNYGSSITEATHLAPS 408
Db 359 QFVPHLSSITLDENVSSGGDIYIGTQWTSPPSDSGGANTFNFKIPDYGSSLAESAQLAPA 418
QY 409 VTPPGGEVLVFMSPKIPGP--GAYSL-PCLLPQEIYISHLASEQAQPTVGEAALLHYVDP 464
```



```
Db 121 GKVIICVPPGQSRSLTSLAQATLPHVIADVRTLDPEVEPLEDVRNVLXHNND-TQPTM 179
QY 181 RLVCMLYTPLTRGGTG--DSFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPL 238
Db 180 RLCCMLYTPLTRGGASGGTDSFVAGRVLTCPGPDNLFVLPPTVEQKTRPFTVPNPL 239
QY 239 SSLNSRAPLPTSGMGISPDNVQSQVFQNGRCTLGRLVGTTPVSLSHVAKIRG--TSNG 296
Db 240 KYLSNSRIPNPIEGMSLSLSPDQTNVQFQNGRCTLGRLVGTTPVSVSLCKFRGRTSGQ 299
QY 297 TVINLTGDTGTFHFFEGPAPICFPDLGGCDWHINMTQFGHSSQTQYD-----VDTPD 350
Db 300 KYLNLTLGDSFMAFAAPAGFPDLGSCDWHIEMSKIPNSS-TQNNPIVNSVKPNSQ 358
QY 351 TFVPHLGGSI--QANGIGSGNYIGVLSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 408
Db 359 QFVPHLPSITLDENVLSGGDYIGTIQWTSPPSDSGANTFWKIPDYGSSLAESQLAPA 418
QY 409 VYPPGFEVLVFFMSKIPGP---GAYSL-PCLLPOEYISHLASEQAPTVEAALLHYVDP 464
Db 419 VYPPGFEVIVFMAISIPGNSGSPNLVPCLLPQEYITHTFISEQAPITQGEAALLHYVDP 478
QY 465 DTGRTLGEFKAYPDGFLTCVPNGASSGPQQLPINGVVFVSWVSRYQLKPVGTASSARG 524
Db 479 DTNRNLGEFKLYPGGYLTCVPNSSSTGPQQLPDGVFVFASVWSRYQLKPVGTGPARG 538
QY 525 RLGLRR 530
Db 539 RLGVRR 544
```

## RESULT 10

```
Q918D1 PRELIMINARY; PRT; 544 AA.
ID Q918D1
AC Q918D1
DT 01-DEC-2001 (T-REMBLrel. 19, Created)
DT 01-DEC-2001 (T-REMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (T-REMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/New Orleans/266/1993/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171834;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/266/1993/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/266/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/266/1993/US;
RA Ando T., Sato Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414402; AAL12959.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 59354 MW; 50387E99852E801D CRC64;
```

Query Match 73.2%; Score 2063; DB 12; Length 544;  
Best Local Similarity 72.0%; Pred. No. 8.3e-160;

```
Matches 393; Conservative 50; Mismatches 85; Indels 18; Gaps 8;
QY 1 MMASKDATSVDCASGAGLVPEVNASDPLAMDVPAGSSATATAGQVNPIDPWIINNF 60
Db 1 MMASKDATSVDCASGAGLVPEVNASDPLAMDVPAGSSATATAGQVNPIDPWIINNF 60
QY 61 VQAPQGEFTTSPNNTPGCVLFDLSGLPHLNPFLHLHSQMYNGWGNVRIMLAGNAFTA 120
Db 61 VQAPQGEFTTSPNNTPGCVLFDLSGLPHLNPFLHLHSQMYNGWGNVRIMLAGNAFTA 120
QY 121 GKIVISCIPIPGFSGHNTIAQAATLPHVIADVRTLDPEVEPLEDVRNVLXHNND-TQPTM 180
Db 121 GKIVISCIPIPGFSGHNTIAQAATLPHVIADVRTLDPEVEPLEDVRNVLXHNND-TQPTM 180
QY 181 RLVCMLYTPLTRGGTG--DSFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPL 238
Db 180 RLVCMLYTPLTRGGTG--DSFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPL 238
QY 239 SSLNSRAPLPTSGMGISPDNVQSQVFQNGRCTLGRLVGTTPVSLSHVAKIRG--TSNG 296
Db 239 SSLNSRAPLPTSGMGISPDNVQSQVFQNGRCTLGRLVGTTPVSLSHVAKIRG--TSNG 296
QY 297 TVINLTGDTGTFHFFEGPAPICFPDLGGCDWHINMTQFGHSSQTQYD-----VDTPD 350
Db 297 TVINLTGDTGTFHFFEGPAPICFPDLGGCDWHINMTQFGHSSQTQYD-----VDTPD 350
QY 300 RALNLTGDSFMAFAAPAGFPDLGSCDWHIEMSKIPNSS-TQNNPIVNSVKPNSQ 358
Db 300 RALNLTGDSFMAFAAPAGFPDLGSCDWHIEMSKIPNSS-TQNNPIVNSVKPNSQ 358
QY 351 TFVPHLGGSI--QANGIGSGNYIGVLSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 408
Db 351 TFVPHLGGSI--QANGIGSGNYIGVLSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 408
QY 359 QFVPHLPSITLDENVLSGGDYIGTIQWTSPPSDSGANTFWKIPDYGSSLAESQLAPA 418
Db 359 QFVPHLPSITLDENVLSGGDYIGTIQWTSPPSDSGANTFWKIPDYGSSLAESQLAPA 418
QY 409 VYPPGFEVLVFFMSKIPGP---GAYSL-PCLLPOEYISHLASEQAPTVEAALLHYVDP 464
Db 409 VYPPGFEVLVFFMSKIPGP---GAYSL-PCLLPOEYISHLASEQAPTVEAALLHYVDP 464
QY 419 VYPPGFEVIVFMAISIPGNSGSPNLVPCLLPQEYITHTFISEQAPITQGEAALLHYVDP 478
Db 419 VYPPGFEVIVFMAISIPGNSGSPNLVPCLLPQEYITHTFISEQAPITQGEAALLHYVDP 478
QY 465 DTGRTLGEFKAYPDGFLTCVPNGASSGPQQLPINGVVFVSWVSRYQLKPVGTASSARG 524
Db 465 DTGRTLGEFKAYPDGFLTCVPNGASSGPQQLPINGVVFVSWVSRYQLKPVGTASSARG 524
QY 479 DTNRNLGEFKLYPGGYLTCVPNSSSTGPQQLPDGVFVFASVWSRYQLKPVGTGPARG 538
Db 479 DTNRNLGEFKLYPGGYLTCVPNSSSTGPQQLPDGVFVFASVWSRYQLKPVGTGPARG 538
QY 525 RLGLRR 530
Db 539 RLGVRR 544
```

## RESULT 11

```
Q9IV41 PRELIMINARY; PRT; 544 AA.
ID Q9IV41
AC Q9IV41
DT 01-OCT-2000 (T-REMBLrel. 15, Created)
DT 01-OCT-2000 (T-REMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (T-REMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Valetta/95/Malta.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122920;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Valetta/95/Malta;
RX MEDLINE=20404983; PubMed=10949950;
RA Green J., Valls J., Gallimore C.I., Koopmans M., Hale A.D.,
Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2004)
DR EMBL; AJ277616; CAB89097.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58415 MW; B6741846BD82B6D1 CRC64;
```

Query Match 72.6%; Score 2046; DB 12; Length 544;  
Best Local Similarity 71.6%; Pred. No. 2e-158;  
Matches 391; Conservative 53; Mismatches 84; Indels 18; Gaps 8;



QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSSTAVATAGVNPIDPWIINN 60  
DB 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSSTAVATAGVNPIDPWIINN 60  
QY 61 VQAPQGEFTISNNTPGGVLFDSLGPHLNPFLHLSOMYNGWGNMVRVIRIAGNAFTA 120  
DB 61 VQAPQGEFTISNNTPGGVLFDSLGPHLNPFLHLSOMYNGWGNMVRVIRIAGNAFTA 120  
QY 121 GKIVSCIPPGSGSHNLTIATLPPHVIADVRTLDPIEVPLEDVRNVLPHNDRNQPTM 180  
DB 121 GKIVSCIPPGSGSHNLTIATLPPHVIADVRTLDPIEVPLEDVRNVLPHNDRNQPTM 180  
QY 181 RLVCMLYPLRTGGGTG--DSFVWAGRVMTCPSPDNFLVPPVVEQKTRFTPLNPL 238  
DB 181 RLVCMLYPLRTGGGTG--DSFVWAGRVMTCPSPDNFLVPPVVEQKTRFTPLNPL 238  
QY 239 SSLNSRAPLPIGSGHISPDNVQSFQNGRCITDGLRLVGTTPVSLSHVAKIRG--TSNG 296  
DB 239 SSLNSRAPLPIGSGHISPDNVQSFQNGRCITDGLRLVGTTPVSLSHVAKIRG--TSNG 296  
QY 297 TVINLTDLGTFFHPEGPAPIGFDPDLGGCDWHINMTQFHSSTQYD-----VDTPD 350  
DB 297 TVINLTDLGTFFHPEGPAPIGFDPDLGGCDWHINMTQFHSSTQYD-----VDTPD 350  
QY 351 TPVPHLGSII--QANGIGSGNYIGVLSWVSPPHSGSGVDLWKIPNYGSSITEATHLAPS 408  
DB 351 TPVPHLGSII--QANGIGSGNYIGVLSWVSPPHSGSGVDLWKIPNYGSSITEATHLAPS 408  
QY 359 QVPHLSITLIDENVSSTGDIYCTQWTSPPSDSGGANTFWKIPDYSSLAESQLAPA 418  
DB 359 QVPHLSITLIDENVSSTGDIYCTQWTSPPSDSGGANTFWKIPDYSSLAESQLAPA 418  
QY 409 VYPGGEVLVFMFSKIPGP---GAYSL-PCLLPOEYISHLASEOAPTVEAALLHYVDP 464  
DB 409 VYPGGEVLVFMFSKIPGP---GAYSL-PCLLPOEYISHLASEOAPTVEAALLHYVDP 464  
QY 419 VYPLVNEVIVFMACIRGPNCGSNLVPCLSPQOEYITHFISEQAPQGEAALLHYVDP 478  
DB 419 VYPLVNEVIVFMACIRGPNCGSNLVPCLSPQOEYITHFISEQAPQGEAALLHYVDP 478  
QY 465 DTGRTLGFEKAYPDGFLTCVNGASSGPOQLPVGIVFVSVSVFYLKPKVGTASSARG 524  
DB 465 DTGRTLGFEKAYPDGFLTCVNGASSGPOQLPVGIVFVSVSVFYLKPKVGTASSARG 524  
QY 479 DTNRNLGEFLKYPGGYLTCVNPNSSTGPOQLDGVFVSVSVFYLKPKVGTAGPARG 538  
DB 479 DTNRNLGEFLKYPGGYLTCVNPNSSTGPOQLDGVFVSVSVFYLKPKVGTAGPARG 538  
QY 525 RLGLRR 530  
DB 525 RLGLRR 530  
QY 539 RLGLRR 544  
DB 539 RLGLRR 544

## RESULT 12

QY9S14 PRELIMINARY; PRT; 540 AA.  
ID QY9S14;  
AC QY9S14;  
DT 01-MAY-1999 (TremBLrel. 10, Created)  
DT 01-MAY-1999 (TremBLrel. 10, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Norwalk virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=11983;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Norwalk-like;  
RX MEDLINE=20254531; PubMed=10795514;  
RA Schreier E., Doering F., Kuenkel U.;  
RT "Molecular epidemiology of outbreaks of gastroenteritis associated  
with small round structured viruses in Germany in 1997/98.";  
RL Arch. Virol. 145:443-453 (2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Norwalk-like;  
RA Schreier E., Doering F., Kuenkel U.;  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBSJ databases.  
DR EMBL; AF093797; AAC64603.1;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat\_1.  
SQ SEQUENCE 540 AA; 58018 MW; 647843487A654CBE CRC64;

## Query Match

72.5%; Score 2044.5; DB 12; Length 540;

Best Local Similarity 71.2%; Pred. No. 2.7e-158;  
Matches 386; Conservative 59; Mismatches 82; Indels 15; Gaps 7;  
QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSSTAVATAGVNPIDPWIINN 60  
DB 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSSTAVATAGVNPIDPWIINN 60  
QY 61 VQAPQGEFTISNNTPGGVLFDSLGPHLNPFLHLSOMYNGWGNMVRVIRIAGNAFTA 120  
DB 61 VQAPQGEFTISNNTPGGVLFDSLGPHLNPFLHLSOMYNGWGNMVRVIRIAGNAFTA 120  
QY 121 GKIVSCIPPGSGSHNLTIATLPPHVIADVRTLDPIEVPLEDVRNVLPHNDRNQPTM 180  
DB 121 GKIVSCIPPGSGSHNLTIATLPPHVIADVRTLDPIEVPLEDVRNVLPHNDRNQPTM 180  
QY 181 RLVCMLYPLRTGGGTG--DSFVWAGRVMTCPSPDNFLVPPVVEQKTRFTPLNPL 238  
DB 181 RLVCMLYPLRTGGGTG--DSFVWAGRVMTCPSPDNFLVPPVVEQKTRFTPLNPL 238  
QY 239 SSLNSRAPLPIGSGHISPDNVQSFQNGRCITDGLRLVGTTPVSLSHVAKIRG--TSNG 296  
DB 239 SSLNSRAPLPIGSGHISPDNVQSFQNGRCITDGLRLVGTTPVSLSHVAKIRG--TSNG 296  
QY 297 TVINLTDLGTFFHPEGPAPIGFDPDLGGCDWHINMTQFHSSTQYD-----TTPTD 354  
DB 297 TVINLTDLGTFFHPEGPAPIGFDPDLGGCDWHINMTQFHSSTQYD-----TTPTD 354  
QY 301 RVLNLSYDGTGTPFVPLESPAPVGPDPDGGCDWHVGTFFEARDDQDPQSNVTFATNDSSFPV 360  
DB 301 RVLNLSYDGTGTPFVPLESPAPVGPDPDGGCDWHVGTFFEARDDQDPQSNVTFATNDSSFPV 360  
QY 355 HLGSII--QANGIGSGNYIGVLSWVSPPHSGSGVDLWKIPNYGSSITEATHLAPSVPV 411  
DB 355 HLGSII--QANGIGSGNYIGVLSWVSPPHSGSGVDLWKIPNYGSSITEATHLAPSVPV 411  
QY 412 PGFGEVLVFMFSKIPGP---AYSPLCLPOEYISHLASEOAPTVEAALLHYVDPDTCR 468  
DB 412 PGFGEVLVFMFSKIPGP---AYSPLCLPOEYISHLASEOAPTVEAALLHYVDPDTCR 468  
QY 419 PGFGEVLVFMFSKIPGP---AYSPLCLPOEYISHLASEOAPTVEAALLHYVDPDTCR 478  
DB 419 PGFGEVLVFMFSKIPGP---AYSPLCLPOEYISHLASEOAPTVEAALLHYVDPDTCR 478  
QY 469 TLGEFKAYPDGFLTCVNGASSGPOQLPVGIVFVSVSVFYLKPKVGTASSARGRLGL 528  
DB 469 TLGEFKAYPDGFLTCVNGASSGPOQLPVGIVFVSVSVFYLKPKVGTASSARGRLGL 528  
QY 479 NLGEFLKYPGGYLTCVNPNSSTGPOQLPVGIVFVSVSVFYLKPKVGTASSARGRLGL 537  
DB 479 NLGEFLKYPGGYLTCVNPNSSTGPOQLPVGIVFVSVSVFYLKPKVGTASSARGRLGL 537  
QY 529 RR 530  
DB 529 RR 530  
QY 538 RR 539  
DB 538 RR 539  
RESULT 13  
QY9V36 PRELIMINARY; PRT; 544 AA.  
ID QY9V36;  
AC QY9V36;  
DT 01-OCT-2000 (TremBLrel. 15, Created)  
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus HU/NLV/Thistlehall/90/UK.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=122925;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HU/NLV/Thistlehall/90/UK;  
RX MEDLINE=20404883; PubMed=10949950;  
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,  
RT Clegg J.C., Chamberlain J., Brown D.W.G.;  
RL "Capsid protein: diversity among 'Norwalk-like' viruses.";  
RL Virus Genes 20:227-236 (2000).  
DR EMBL; AJ277621; CAB89102.1;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat\_1.  
SQ SEQUENCE 544 AA; 58676 MW; 4D261DFA42621EDE CRC64;

## Query Match

72.5%; Score 2044; DB 12; Length 544;

Best Local Similarity 71.6%; Pred. No. 3e-158;  
Matches 391; Conservative 53; Mismatches 84; Indels 19; Gaps 8;

```
QY 1 MMASKATSSVDGASGAGQLVPEVNASDPLAMPDVAGSSSTAVATAGQVNPIDPHIINF 60
Db 1 MMASKADATPSADGATGAGQLVPEVNTADIPIDVAGSSSTALATAGQVNLIDPHIINF 60
QY 61 VOAPQGEFTISNNPTGGVLFDLISGLPHLNPFLHLSQMYNGWGNMRVRIMLAGNAFTA 120
Db 61 VOAPQGEFTISNNPTGGVLFDLISGLPHLNPFLHLSQMYNGWGNMRVRVVLAGNAFTA 120
QY 121 GKIIIVSCIPPGGSHNLTAQATLPHVIADVRTLDPIEVPLEDRVNLPHNDRNQOTM 180
Db 121 GKVIICVPPGQSRSLTAQATLPHVIADVRTLDPIEVPLEDRVNLPHNDRNQOTM 179
QY 181 RLVCMLYTPRTGGGTG--DSFVAGRVMTCPSPDFNLFVLPPTVEQKTRFTPLNPL 238
Db 180 RLICMLYTPRTGGGTGDSFVAGRVMTCPSPDFNLFVLPPTVEQKTRFTPLNPL 239
QY 239 SSLSNSRAPLPIISGMGISPNDVQVQFONGRCTLDGRLVGTTPVSLSHVAKIRG--TSNG 296
Db 240 KYLSNSRIENPIEGMSLSPDQTNQVQFONGRCTLDGRLVGTTPVSLSHVAKIRG--TSNG 299
QY 297 TVINTELDTPPHPEGPAPIGFDPDLGGCDWHINMTQFGHSSQTOYD-----VDTPPD 350
Db 300 KVLNLTELDSSPMAFAAPAGFPDLGSCDWHIEMSKIPNST-TQNNPIVTDVSKPNSQ 358
QY 351 TFVPHLGGSI--QANGIGSGNYIGVLSWVSPSPHSPSGQVDLWKIPNYGSSITEATHLAPS 408
Db 359 QFVPHLSSITLDENVLSGGDYGNIQWTLVLLVILXEPKTNFKWIPDYGSSLAESQLAPA 418
QY 409 VYPPGGEVLVFMSPKIPGP--GAYSL-PCLLPQEIYISHLASQAPTVEAALLHYVDP 464
Db 419 VYPPGNEVIVFMASIPGNQSGSNLPVCLLPQEIYITHFISEQAPIQGEAALLHYVDP 478
QY 465 DTGRTLGEKAYPDGFLTCVPGNASSGQQLPINGVVFVSWVSRFYQLKPVGTASSARG 524
Db 479 DTNRNLGEFLKYPGYLTCVPNSSSTGPQQLPDGVFVSWVSRFYQLKPVGTAGPARG 538
QY 525 RLGLRR 530
Db 539 RLGVRR 544
```

```
RESULT 14
Q8V7J5 PRELIMINARY; PRT; 531 AA.
AC Q8V7J5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein (Fragment).
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RA Yoda T., Terano Y., Suzuki Y., Yamazaki K., Oishi I., Kuzuguchi T.,
RA Kawamoto H., Utagawa E., Takino K., Shibata T.;
RT "Characterization of Norwalk virus G1 Specific Monoclonal Antibodies
RT Generated against Escherichia coli Expressed Capsid Protein and
RT Recombinant Fragments.", the EMBL/GenBank/DBJ databases.
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB063325; BAB83500.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
FT NON_TER 531
FT NON_TER 531
SQ SEQUENCE 531 AA; 56747 MW; 9688132FDDCFB88F CRC64;

Query Match 72.0%; Score 2031; DB 12; Length 531;
Best Local Similarity 72.2%; Pred. No. 3.3e-157;
```

```
Matches 385; Conservative 50; Mismatches 80; Indels 18; Gaps 8;
QY 8 ATSSVDGASGAGQLVPEVNASDPLAMPDVAGSSSTAVATAGQVNPIDPHIINFVQAPQGE 67
Db 1 ATPSADGATGAGQLVPEVNTADIPIDVAGSSSTALATAGQVNLIDPHIINFVQAPQGE 60
QY 68 FTISPNNPTGGVLFDLISGLPHLNPFLHLSQMYNGWGNMRVRIMLAGNAFTAAGKIIVSC 127
Db 61 FTISPNNPTGGVLFDLISGLPHLNPFLHLSQMYNGWGNMRVRVVLAGNAFTAAGKVIIC 120
QY 128 IPPGFGSHNLTAQATLPHVIADVRTLDPIEVPLEDRVNLPHNDRNQOTMRLVLCMLY 187
Db 121 VPPGQSRSLTAQATLPHVIADVRTLDPIEVPLEDRVNLPHNDRNQOTMRLVLCMLY 179
QY 188 TPLRTGGGTG--DSFVAGRVMTCPSPDFNLFVLPPTVEQKTRFTPLNPLSSLSNSR 245
Db 180 TPLRTGASGTTDSFVAGRVMTCPSPDFNLFVLPPTVEQKTRFTPLNPLKYLNSR 239
QY 246 APLPISGMGISPNDVQVQFONGRCTLDGRLVGTTPVSLSHVAKIRG--TSNGTVINILTE 303
Db 240 IPNPIEGMSLSPDQTNQVQFONGRCTLDGRLVGTTPVSLSHVAKIRG--TSNGTVINILTE 299
QY 304 LDGTPPHPEGPAPIGFDPDLGGCDWHINMTQFGHSSQTOYD-----VDTPPD 357
Db 300 LDGSPMAFAAPAGFPDLGSCDWHIEMSKIPNST-TQNNPIVTVSWVSKNSQGFVPHLS 358
QY 358 SI--QANGIGSGNYIGVLSWVSPSPHSPSGQVDLWKIPNYGSSITEATHLAPSYPGFG 415
Db 359 SITLDDNVSSGGDYGNIQWTLVLLVILXEPKTNFKWIPDYGSSLAESQLAPA 418
QY 416 EVLVFMSPKIPGP--GAYSL-PCLLPQEIYISHLASQAPTVEAALLHYVDPDPTGRTL 471
Db 419 EVLVFMASIPGNQSGSNLPVCLLPQEIYITHFISEQAPIQGEAALLHYVDPDPTNRNL 478
QY 472 EFKAYPDGFLTCVPGNASSGQQLPINGVVFVSWVSRFYQLKPVGTASSARG 524
Db 479 EFKLYPGGYLTCVPNSSSTGPQQLPDGVFVSWVSRFYQLKPVGTAGPARG 531
```

```
RESULT 15
Q8UW44 PRELIMINARY; PRT; 546 AA.
AC Q8UW44;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RA Katayama K., Takeda N., Natori K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Natori K., Takeda N.;
RT "Genetic and antigenic relationship among Norwalk-like viruses.",
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB078335; BAC05516.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 546 AA; 58834 MW; 1965F054E2C481E6 CRC64;

Query Match 72.0%; Score 2029.5; DB 12; Length 546;
Best Local Similarity 69.4%; Pred. No. 4.6e-157;
Matches 379; Conservative 68; Mismatches 82; Indels 17; Gaps 6;
```

[illegible]

Search completed: June 1, 2004, 13:53:10  
Job time : 31.8574 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.5352 Seconds  
(without alignments)  
4985.230 Million cell updates/sec

Title: US-09-926-799-3  
Perfect score: 2895  
Sequence: 1 MMASKADPQADGASGAGQ.....QLKPVGTASTARSLGVRRRI 546

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2887	99.7	546	2 B37491	major capsid prote
2	2030.5	70.1	530	2 B37471	capsid protein - N
3	1149.5	39.7	542	2 S60616	capsid protein - h
4	1111	38.4	539	2 S40111	capsid protein - E
5	341.5	11.8	576	2 A53982	capsid protein - h
6	325	11.2	2344	1 RRWRH	genome polyprotein
7	321	11.1	2344	2 S64740	genome polyprotein
8	318	11.0	2344	2 S55399	genome polyprotein
9	299	10.3	703	1 C48562	coat protein - San
10	293.5	10.1	702	1 A48562	coat protein - San
11	284	9.8	668	1 VCMWFC	coat protein - fel
12	275.5	9.5	668	2 JQ2354	capsid protein - f
13	275.5	9.5	668	2 JQ2356	capsid protein - f
14	268.5	9.3	671	1 VCMWF9	coat protein - fel
15	267	9.2	668	1 VCMWFF	coat protein - fel
16	145.5	5.0	2194	1 GNNYE7	genome polyprotein
17	145	5.0	2206	2 S03822	genome polyprotein
18	143	4.9	2206	1 GNNY4P	genome polyprotein
19	140.5	4.9	2206	1 GNNY27	genome polyprotein
20	139	4.8	3473	1 A46112	genome polyprotein
21	139	4.8	3473	2 S27927	polyprotein - rice
22	131.5	4.5	2205	1 GNNY2W	genome polyprotein
23	128.5	4.4	2207	1 GNNY5P	genome polyprotein
24	127.5	4.4	3085	2 T00327	polyprotein - infe
25	123	4.2	2175	1 GNNYBE	genome polyprotein
26	122.5	4.2	2209	1 GNNY3P	genome polyprotein
27	122.5	4.2	13055	2 T16580	hypothetical prote
28	121.5	4.2	2209	1 GNNY2P	genome polyprotein
29	121	4.2	1742	2 S76110	hypothetical prote

30	119	4.1	2164	1 GNNY89	genome polyprotein
31	119	4.1	2185	1 GNNYSV	genome polyprotein
32	119	4.1	2185	1 GNNYSV	genome polyprotein
33	117.5	4.1	2207	1 GNNY1P	genome polyprotein
34	117.5	4.1	3972	2 S75251	hypothetical prote
35	117	4.0	2207	2 S09553	genome polyprotein
36	116.5	4.0	2150	1 GNNYH2	genome polyprotein
37	116	4.0	1200	2 T17404	hyalin - sea urchi
38	116	4.0	2185	1 JQ2021	genome polyprotein
39	115	4.0	888	1 GNLJHD	pol polyprotein -
40	113.5	3.9	2333	1 GNNY2F	genome polyprotein
41	113	3.9	896	2 I56563	interleukin-3 rece
42	112.5	3.9	708	1 M2XRLL	structural protein
43	112.5	3.9	2214	1 A48548	genome polyprotein
44	112	3.9	2179	1 GNNYH4	genome polyprotein
45	111.5	3.9	1344	2 T14316	rig-1 protein - mo

ALIGNMENTS

RESULT 1

B37491 major capsid protein [similarity] - Southampton virus

N:Alternate names: orf2 protein

C:Species: Southampton virus

C:Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 18-Aug-2000

C:Accession: B37491

RefSeq: P. R. J. Gaul, E. O. Ashley, C. R. Clarke, I. M. R.

Sequence 259,546,519,193

A:Title: Sequence and genome organization of a human small round-structured (Norwalk-like

A:Reference number: A37491; MUID:93142023; PMID:8380940

A:Accession: B37491

A:Status: not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-546 <LAN>

A:Cross-references: GB:L07418; NID:gl236787; PIDN:AAA92984.1; PID:g295114

A>Note: sequence extracted from NCBI backbone (NCBIP:123458)

A>Note: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype

C:Superfamily: human calicivirus capsid protein

C:Keywords: glycoprotein

F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 99.7%; Score 2887; DB 2; Length 546;

Best Local Similarity 99.6%; Pred. No. 2.9e-207;

Matches 544; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY	1	MMASKADPQADGASGAGQLVPEVNTADPLPMEPVAGTTTAVATAGQVNMIDPWVNNF	60
DB	1	MMASKADPQADGASGAGQLVPEVNTADPLPMEPVAGTTTAVATAGQVNMIDPWVNNF	60
QY	61	VQSPQGETTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGMVRVILLAGNAFSA	120
DB	61	VQSPQGETTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGMVRVILLAGNAFSA	120
QY	121	GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLTEPIEMPLEDVRNVLHYHTNDNQPMR	180
DB	121	GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLTEPIEMPLEDVRNVLHYHTNDNQPMR	180
QY	181	LVCMLYTLRTGGGSGNSDSFVAGRVLTAPSSDFSFLFLVPPPTIEQKTRAFVNIPLQ	240
DB	181	LVCMLYTLRTGGGSGNSDSFVAGRVLTAPSSDFSFLFLVPPPTIEQKTRAFVNIPLQ	240
QY	241	TLNSRRFSLTQGMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFVRGKKNQGAR	300
DB	241	TLNSRRFSLTQGMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFVRGKKNQGAR	300
QY	301	TLNLTVDGKPFMAFDSAPVGFDFGKCDHMRISKTPNNNTSSGDPMRSVSQTNVQGF	360
DB	301	TLNLTVDGKPFMAFDSAPVGFDFGKCDHMRISKTPNNNTSSGDPMRSVSQTNVQGF	360
QY	361	VPHLGSIQFDEVFNPTGDYGTTEWISQPSPTPGTDNLNWEIPDYGSSLSQAANLAPPV	420
DB	361	VPHLGSIQFDEVFNPTGDYGTTEWISQPSPTPGTDNLNWEIPDYGSSLSQAANLAPPV	420

```

Db      361 VPHLSIQFDEVFNHPTGDYITIEWISQSPSTPPTDINLWEIPDYGSSLSQAANLAPPV 420
Qy      421 PPGGEGEALVYFVSAPFPNNRSAPNDVPCLLPOEYITHFVSEQAPTMGDAALHYVDPD 480
Db      421 PPGGEGEALVYFVSAPFPNNRSAPNDVPCLLPOEYITHFVSEQAPTMGDAALHYVDPD 480
Qy      481 TNRNLGEFKLYPGGYLTCPVNGVAGPQQLPLNGVFLFVSVWSRFPYQKPVGTASTARS 540
Db      481 TNRNLGEFKLYPGGYLTCPVNGVAGPQQLPLNGVFLFVSVWSRFPYQKPVGTASTARG 540
Qy      541 LGVRR 546
Db      541 LGVRR 546

RESULT 2
B37471
capsid protein - Norwalk virus
C:Species: Norwalk virus
C>Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000
C:Accession: B37471
R:Jiang, X.; Wang, M.; Wang, K.; Estee, M.K.
Virology 195, 51-61, 1993
A:Title: Sequence and genomic organization of Norwalk virus.
A:Reference number: A37471; MUID:93303939; PMID:8391187
A:Accession: B37471
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: genomic RNA
A:Residues: 1-530 <JIA>
A:Cross-references: GB:M87661; NID:G1061311; PIDN:AAB50466.1; PID:G1061313
A:Note: sequence extracted from NCBI backbone (NCBI:P134157)
C:Superfamily: human calicivirus capsid protein

Query Match 70.1%; Score 2030.5; DB 2; Length 530;
Best Local Similarity 69.2%; Pred. No. 1.8e-143;
Matches 378; Conservative 70; Mismatches 81; Indels 17; Gaps 6;

Qy      1 MMASKADAPQASDAGSAGAGOLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWVNNF 60
Db      1 MMASKADATSSVDCASGAGOLVPEVNASDPLADPVGASTAVATAGQVNPIDPWVNNF 60
Qy      61 VQSPQGEFTTSPNNTPGDILFDLQGLPHLPFLSHLSQMYNGVGNMVRVILLAGNAFSA 120
Db      61 VQAPQGEFTTSPNNTPGDVLFDLSGLPHLPFLSHLSQMYNGVGNMVRVILLAGNAFTA 120
Qy      121 GKTIIVCCVPPGFTSSSLTIAQATLFPHVIAVDRVLEPIEMPLEDVRNLVHTND-NQPTM 179
Db      121 GKTIIVCCIPPGFGSHNLTIAQATLFPHVIAVDRVLDPIEVLDPEDVRNLVPHNNDRNQPTM 180
Qy      180 RLVCMLYPTLPTGGSGNSDSFVAGVLTAPSDSFSLFVLPPTIEQKTRAFVTPNIPL 239
Db      181 RLVCMLYPTLPTGGGTG--DSFVAGVMTCPSPDFNPLFLVPTVEQKTRPPTLNLPL 238
Qy      240 QTLNSRFPSSLIOGMILSPDASQVVFQNGRCLLDGQLGTTTATSPATSQQLFRVRGKINQGA 299
Db      239 SSLNSRAPLPIISIGISPDNVQSVQFQNGRCLLDGLVGTTPVSLSHVAKIRGTSN--G 296
Qy      300 RTLNLTEVDGKPFMAFSPAPVGPDPFGKCDWHMRISKTPNNTSSGDPMSVSQVTVNQV 359
Db      297 TVINLTEDGTTPFPFEPGAPIGPFDLGGCDWHINMTQFGHSSQT-----QYDVTDPDT 351
Qy      360 FVPHLGSTQFDEVFNHPTGDYITIEWISQSPSTPPTDINLWEIPDYGSSLSQAANLAPP 419
Db      352 FVPHLGSIQANGI-----GSGNVGVLSWISPSHSPSGSQVDLWKLPNYGSSITEATHLAPS 408
Qy      420 VPPPGFGEALVYFVSAPFPNNRSAPNDVPCLLPOEYITHFVSEQAPTMGDAALHYVD 479
Db      409 VYPPGFGVGVVFFMSKMPG---GAYNLPCLLPQEIYISHLASEQAPTVGEAALLHYVD 464
Qy      480 DTNRNLGEFKLYPGGYLTCPVNGVAGPQQLPLNGVFLFVSVWSRFPYQKPVGTASTARS 539
Db      465 DTGRNLGEFKAYDPDGLTFCVPNGASSGPPQIPNGVFFVSVWSRFPYQKPVGTASSARG 524

```

```

Qy      540 RLGVRR 545
Db      525 RLGLRR 530

RESULT 3
S60616
capsid protein - human calicivirus (strain Melksham)
C:Species: human calicivirus
A:Variety: strain Melksham
C>Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 28-Jul-2000
C:Accession: S60616
R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
Virus Res 37, 271-283, 1995
A:Title: Capsid diversity in small round-structured viruses: molecular characterization
A:Reference number: S60615; MUID:96136658; PMID:8533462
A:Accession: S60616
A:Molecule type: genomic RNA
A:Residues: 1-542 <GRE>
A:Cross-references: EMBL:X81879; NID:G976077; PIDN:CAA57462.1; PID:G976079
A:Experimental source: strain Melksham
A:Note: it is uncertain whether Met-1 or Met-3 is the initiator
C:Superfamily: human calicivirus capsid protein
C:Keywords: capsid protein; coat protein

Query Match 39.7%; Score 1149.5; DB 2; Length 542;
Best Local Similarity 44.7%; Pred. No. 8.4e-78;
Matches 256; Conservative 82; Mismatches 176; Indels 59; Gaps 19;

Qy      1 MMASKADAPQASDAGSAGAGOLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWVNNF 60
Db      1 MMASKADAPSTDGAAG---LVPESNN-EVVALPEVAGAAALAPVTGQTNIIDPWIRANF 56
Qy      61 VQSPQGEFTTSPNNTPGDILFDLQGLPHLPFLSHLSQMYNGVGNMVRVILLAGNAFSA 120
Db      57 VQAPNGEFTVSPNAPQEVLLNLELPELNPYLALHARMYNGVAGMEVQVMLAGNAFTA 116
Qy      121 GKTIIVCCVPPGFTSSSLTIAQATLFPHVIAVDRVLEPIEMPLEDVRNLVHTND-NQPTM 179
Db      117 GKLVFAAVPPHPFVENLSPQKITMFPHVIIDVLTLEPVLPLPDVRNSFFHYNQKDDPKM 176
Qy      180 RLVCMLYPTLPTGGSGNSDSFVAGVLTAPSDSFSLFVLPPTIEQKTRAFVTPNIPL 239
Db      177 RIVANLYTPLSR-NGSGD-DVFTVSCRVLTRPSPDFDTYLVPTVESKTKPFTLPILT 234
Qy      240 QTLNSRFPSSLIOGMILSPDASQVVFQNGRCLLDGQLGTTTATSPATSQQLFRVRGKI---- 295
Db      235 GELNSRFPVPIDQMYTSPNEVISVQCNGRCLTDGLQGTTLQVSGICAFKGEVTAHL 294
Qy      296 --NQARTLNLTEDVGKPF-MAFDSAPVGPDPF-GKCDWHMRISKTPNNTSSGDPM--- 348
Db      295 HDNDHLNVTITNLNGSPFDPSEDIPAPLGVDPFQGRVGVISORDKQNAAGHSEPANRG 354
Qy      349 RSVSVQTVNQGVFVPHLGSIQFDEVFNHPTGDYITIEW-----ISQP--STPPTD--- 397
Db      355 HDVAVPTTYATYTPKLGQIQ-----IGT--WQTDLTWQVQKFTPVGLNDTE 400
Qy      398 -INLWEIPDYGSSLSQAANLAP---VPPPGFGEALVYFVSAPFPNNRSAPNDVPCLLP 453
Db      401 HFNQWVPRYACALNLTNLAPSAVAPFP---GEFLFRSHLPLKGGYGNP-AIDCLLP 456
Qy      454 QEYITHFVSEQAPTMGDAALHYVDPTNRNLGEFKLYPGGYLTCPVNGVAGPQQLPLN 513
Db      457 QEWVQHFYQEAAPSSEVALVRIYINPDTGRALFEAKLHRAGMTVSSN--TSAPVVVPAN 514
Qy      514 GVFLFVSVWSRFPYQKPVGTASTARSRLGVRR 546
Db      515 GYFRFDSWNQFYSLAPMGTGN-----GRRRV 541

RESULT 4
S40111
capsid protein - human calicivirus (strain Bristol isolate B493)

```

C:Species: human calicivirus  
A:Variety: strain Bristol isolate B493  
C:Date: 25-Dec-1994 #sequence\_revision 27-Feb-1997 #text\_change 28-Jul-2000  
C:Accession: S40111  
R:Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
submitted to the EMBL Data Library, December 1993  
A:Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de  
A:Reference number: S40111  
A:Accession: S40111  
A:Molecule type: genomic RNA  
A:Residues: 1-539 <GRE>  
A:Cross-references: EMBL:X76716; NID:G436410; PIDN:CAAS4134.1; PID:G436411  
A:Experimental source: human enteric calicivirus strain Bristol isolate B493  
C:Superfamily: human calicivirus capsid protein  
C:Keywords: capsid protein; coat protein

Query Match 38.4%; Score 1111; DB 2; Length 539;  
Best Local Similarity 44.3%; Pred. No. 6.1e-75;  
Matches 250; Conservative 76; Mismatches 190; Indels 48; Gaps 17;

QY 1 MMWASKADPQSDAGSAGAGQVPEVNTADPLPMEPVAGTTAVATAGQVNMIDPMVNNF 60  
DB 1 MKMASNDANPS-DGS--AANLVPEVNN-EVMALEPVGAAIAAPVAGQONVIDPMIRNF 56  
QY 61 VQSPQGEFTISNNTPGDIILFDLQGLPHLNPFLSHLSONYNGWGNRVRIILAGNAPSA 120  
DB 57 VOAPQGEFTVSPNAPGEILWSAPLGPDLNPLYLHSLRMYNGYAGGFEVQVILAGNAFTA 116  
QY 121 GKIIIVCCVPPGTSSTLTAAQATLPHVIADVRLTEPIEMPLEDVNRVLYHTND--NPTM 179  
DB 117 GKVIIFAAVPPNPTEGLSPQVMTMPPHIIIVDRQLEPVLIPLDVNRNPFYHYNQANDSTL 176  
QY 180 RLVCMLYTLPLRTGGSGNSDSFVAGRVLTAPSSDFSLFLVPTIEQKTRAFVTPNIPL 239  
DB 177 KLIALMYTELR--ANNAGDDVTVSCRVLTRESDFDFELFVPTVESRTKFTVPVLTV 234  
QY 240 QTLNSRPSLSIQMILSPDASQVQFQNGRCLIGQLLGGTTPATSGQLFRVRGKINQGA 299  
DB 235 EEMSNRFPLEKLYTGSSAFVQVQNGRCRTDGLVLTGTTQLSAVINCFNRGVDTHIA 294  
QY 300 ----RTLNLTEVDGKPFMAFDS-PAPVGPDPF-GKCDWHMRISKTPNNTSSGD---PMRS 350  
DB 295 GSHDTMNLASQWNSYDTEISAPLGPDPFVGK-----IQGLLTQTTRADGGSTRAHK 348  
QY 351 VSVQTNVQGVFPHLGSIQDFEVNHTPDY-----IGTIEWISQFSTPPGTDINL 400  
DB 349 ATVSTGSHVHTFKLGSVQTTDTNN---DFQAGQNTKFTPVGVIQGDHQNHP---QQ 401  
QY 401 WEIPDYGSSLSQAANLAPVPFPPGGEALVYFVSAPPGNNRSAPN-DVPCLLPQRYITH 459  
DB 402 WLLPNYSGRTHNVHLAPAVPTFGEQQLFFRSTMPGCS--GYPNMNLDCLLPQEWVLH 459  
QY 460 FVSEQAPTMGDAALLHYVDPDNRNLGEPLKYPGGVLTCPVNGVAGPQO--LPLNGVPL 517  
DB 460 FYQEAPASQDVALLRFVNPDPGRVLFCEKLHKSGYITV-----AHTGPVLDLVPNGYPR 515  
QY 518 FVSVWSRFTQKLPVGTASTARSRL 541  
DB 516 FDSWNQFTLAPMGNGTGRRAL 539

RESULT 5  
A53982  
capsid protein - European brown hare syndrome virus  
C:Species: European brown hare syndrome virus  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 28-Jul-2000  
C:Accession: A53982  
R:Wirthlich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Eskens, U.; Haas, B.; Thiel, H.  
J. Virol. 68, 5164-5173, 1994  
A:Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease  
A:Reference number: A53982; MUID:94309183; PMID:7518531  
A:Accession: A53982  
A>Status: preliminary

A:Molecule type: mRNA  
A:Residues: 1-576 <WIR>  
A:Cross-references: GB:093199  
C:Superfamily: human calicivirus capsid protein

Query Match 11.8%; Score 341.5; DB 2; Length 576;  
Best Local Similarity 25.1%; Pred. No. 1.5e-17;  
Matches 142; Conservative 66; Mismatches 217; Indels 141; Gaps 23;

QY 7 DAPOSADGASGAGQLV-----PEVNTADPLPMEPVAGTTTAVATAG-----QVN 50  
DB 8 DAGTATTASVPGTTTDCMDPGVASTDVVTADNV-----AASVATAGIGPPPQQAS 59  
QY 51 MTDPMVNNFVSPQGEFTISNNTPGDIILFDLQGLPHLNPFLSHLSONYNGWGNRVRI 110  
DB 60 POESWRVNFYND---VFTWSVTDAPGSILYSVQHSQNNPFTQVLSQNYAGWAGMQPR 116  
QY 111 ILLAGNASAGKIIVCCVPPGP-TSSSLTIAQATLPHVIADVRLTEPIEMPLEDVNRVL 169  
DB 117 FIVAGSGIFGGLRVCAIIPPGIQIQGLEVRQ---FPHVVIDARSLPEVVTITMPDLRPM 173  
QY 170 YH-TNDNQPTMRLVCMLYTLPLRTGGSGNSDSFVAGRVLTAPSSDFSLFLVPTIEOK 228  
DB 174 YHTGDPGLVPTLVSVYNNLINPEGTTT---AIQVTVETPSEDFEFVLIRAPS--SK 228  
QY 229 TRAFVTPNIPLTL-----SNSRPSLSIQMILSPDASQVQFQNGRCLIGQLLGGTTP 282  
DB 229 TVDSVNPWSMLLTTPVLITGAGSDNRWGPITVGLQPVGGFST---SNRHNWNNGSTYGMSS 285  
QY 283 ATSGQLFRVRGKI-----NOCARTNLTEVDGKPFMAFDSAPVGFDPF 326  
DB 286 PRFDDIDHPSGNVSYPTGSATNWTIETVANAGTATTN-----PISNIAPDGFPM 335  
QY 327 GKCDWHMRISKTPNNTSSGDPMRSVSQTNVQGVFPHLGSIQDFEVNHTPTG-DYIGTIE 385  
DB 336 GAI-----PFGSTTIPTGAW-----VCGQVWNASNGTPYVGTVQ 370  
QY 386 -----WISQPTTPGTDINLWEIPDYGSSLSQAANLAPVPF-----PPGFG 426  
DB 371 AYELGFANGAPSSIRPVTTTGAQLVAKSI--YGVAIAQNOSSAGIIFLSKGMVSTPG-- 426  
QY 427 EALVYFVSAPPGNNRSAPNDVPCLLPQRYITHFVSEQAPTMGDAALLHYVDPDNRNLG 486  
DB 427 ---VAATTYTPQPSAIVTTPGTFVAA-----IGKNTPIMFSAVV-----RRTG 467  
QY 487 EPKLYPGGVLTCPVNGVAGPQOOLPL 512  
DB 468 DVNAGPGS-VNGTQYGVGSQPLSVTL 492

RESULT 6  
RWWRH  
genome polyprotein - rabbit hemorrhagic disease virus  
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C:Species: Rabbit hemorrhagic disease virus  
C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998  
C:Accession: A41039  
R:Meyers, G.; Wirthlich, C.; Thiel, H.J.  
Virology 184, 664-676, 1991  
A:Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing  
A:Reference number: A41039; MUID:91361557; PMID:1840711  
A:Accession: A41039  
A:Molecule type: genomic RNA  
A:Residues: 1-2344 <MEV>  
A:Cross-references: GB:M67473  
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C:Keywords: nucleotidyltransferase

Query Match 11.2%; Score 325; DB 1; Length 2344;  
Best Local Similarity 26.7%; Pred. No. 1.9e-15;  
Matches 118; Conservative 57; Mismatches 183; Indels 84; Gaps 18;

QY 5 SKDAPQSDAGSAGAGQVPEVNTADPLPMEPVAGTTAV-----ATAG-----Q 48

Db 1770 ARAAPQGEAAGTATTASVPGTTTGDG---MDPGVAVTTSVITAENSASIASIATAGIGGPPQQ 1826  
QY 49 VNMIDPWI VNNFVQS PQGEFTTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGVNMR 108  
Db 1827 VDOQETWRTNFYND---VFTWSVADAPGSILYTVQHSFQNNPFTAVLSQMYAGWAGMQ 1883  
QY 109 VRILLAGNAFSAKIIIVCCVPPGF-TSSSLTTIAQATLFPHVADVRLTLEPIEMPLEDVNR 167  
Db 1884 FRFIVAGSVFGRLVRAVIPPGEIIGPGLVRQ---FPHVVIDARSLPEVITITMPDLRP 1940  
QY 168 VLYH-TNDNQPTMRLVCMLYTLRTGGSGNSDSFVAVAGRLTAPSSDFSFLVLP--- 223  
Db 1941 NMYHPTGDPGLVPTLVLSVNNLINFEGGSTS---AIQVTVETRSEDFEFVIRAPSSK 1997  
QY 224 TIEQKTRAFVTPNIPLOTLSN-SRFPSLIQGMILSPDASQVVFQNGRCCLIDGQLLTTP 282  
Db 1998 TVDSISPAGLLTTPVLTVGVGNDRWNGQIVGLQVPVGGFSTC---NRHNLNGSTYGWSS 2054  
QY 283 ATSGQLFRVRGKIN-QGARTLNLTEV-DGKPFMAFDSP---APVGFDPFGKCDWHMRLS 336  
Db 2055 PRFGDIDHRRGSASVSGSNATNLQFWYANAGSAIDNPISQVAPDGFDP-----MS 2105  
QY 337 KTPNNTSSGDPMSVSQVTVNQGVFPHLSIQFDEVNHPGTGDIYGTIEWISQSPSTPGT 396  
Db 2106 FVPFNGPG-----IPAAGWVGFGAI-----WNSNGAPNVT 2136  
QY 397 DINLWEIPDYGSSLSQAANLAP 418  
Db 2137 TVQAYEL---GFATGAPGNLQ 2155  
RESULT 7  
S64740  
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)  
N;Contains: VP60 protein  
C;Species: rabbit hemorrhagic disease virus  
A;Variety: isolate AST/89  
C;Date: 12-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 20-Jun-2000  
C;Accession: S64740; S49018; S46944; S49012  
R;Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.  
submitted to the EMBL Data Library, May 1995  
A;Description: Genomic organization of rabbit hemorrhagic disease virus determined by di  
A;Reference number: S64740  
A;Accession: S64740  
A;Molecule type: genomic RNA  
A;Residues: 1-2344 <CAS>  
A;Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552  
R;Boga, J.; Casais, R.; Marin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, R  
submitted to the EMBL Data Library, July 1993  
A;Description: Molecular cloning, sequence and expression of the capsid protein gene frc  
A;Reference number: S46944  
A;Accession: S46944  
A;Molecule type: genomic RNA  
A;Residues: 1650-2344 <BOG>  
A;Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623  
R;Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.  
Virus Res. 27, 219-228, 1993  
A;Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp  
A;Reference number: S49018; MUID:93255896; PMID:8488721  
A;Accession: S49018  
A;Molecule type: genomic RNA  
A;Residues: 1650-1796 <PAR>  
A;Cross-references: EMBL:Z24757  
A;Experimental source: isolate AST/89  
A;Description: Molecular cloning, sequence and expression of the capsid protein gene frc  
A;Reference number: S65012  
A;Accession: S65012  
A;Molecule type: protein  
A;Residues: 1767-1793;1875-1877, 'X', 1879-1881;1936-1938, 'X', 1940-1941 <PAW>  
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C;Keywords: blocked amino end; polyprotein

Query Match 11.1%; Score 321; DB 2; Length 2344;  
Best Local Similarity 26.7%; Pred. No. 3.8e-15;  
Matches 116; Conservative 55; Mismatches 185; Indels 84; Gaps 18;  
QY 5 SKDAPQADGASGAGQLVPEVNTADPLPMEPVAGPTTAV-----ATAG-----Q 48  
Db 1770 ARTAPQGEAAGTATTASVPGTTTGDG---MDPGVAVTTSVITAENSASIASIATAGIGGPPQQ 1826  
QY 49 VNMIDPWI VNNFVQS PQGEFTTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGVNMR 108  
Db 1827 VDOQETWRTNFYND---VFTWSVADAPGSILYTVQHSFQNNPFTAVLSQMYAGWAGMQ 1883  
QY 109 VRILLAGNAFSAKIIIVCCVPPGF-TSSSLTTIAQATLFPHVADVRLTLEPIEMPLEDVNR 167  
Db 1884 FRFIVAGSVFGRLVRAVIPPGEIIGPGLVRQ---FPHVVIDARSLPEVITITMPDLRP 1940  
QY 168 VLYH-TNDNQPTMRLVCMLYTLRTGGSGNSDSFVAVAGRLTAPSSDFSFLVLP--- 223  
Db 1941 NMYHPTGDPGLVPTLVLSVNNLINFEGGSTS---AIQVTVETRSEDFEFVIRAPSSK 1997  
QY 224 TIEQKTRAFVTPNIPLOTLSN-SRFPSLIQGMILSPDASQVVFQNGRCCLIDGQLLTTP 282  
Db 1998 TVDSISPAGLLTTPVLTVGVGNDRWNGQIVGLQVPVGGFSTC---NRHNLNGSTYGWSS 2054  
QY 283 ATSGQLFRVRGKIN-QGARTLNLTEV-DGKPFMAFDSP---APVGFDPFGKCDWHMRLS 336  
Db 2055 PRFGDIDHRRGSASVSGSNATNLQFWYANAGSAIDNPISQVAPDGFDP-----MS 2105  
QY 337 KTPNNTSSGDPMSVSQVTVNQGVFPHLSIQFDEVNHPGTGDIYGTIEWISQSPSTPGT 396  
Db 2106 FVPFNGPG-----IPAAGWVGFGAI-----WNSNGAPNVT 2136  
QY 397 DINLWEIPDYGSSLSQAANLAP 418  
Db 2137 TVQAYEL---GFATGAPGNLQ 2155  
RESULT 8  
S55399  
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)  
C;Species: rabbit hemorrhagic disease virus  
A;Variety: isolate BS89  
C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 18-Jun-1999  
C;Accession: S55399  
R;Rossi, C.  
submitted to the EMBL Data Library, May 1995  
A;Reference number: S55399  
A;Accession: S55399  
A;Molecule type: genomic RNA  
A;Residues: 1-2344 <ROS>  
A;Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641  
A;Experimental source: isolate BS89  
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C;Keywords: polyprotein  
Query Match 11.0%; Score 318; DB 2; Length 2344;  
Best Local Similarity 26.2%; Pred. No. 6.4e-15;  
Matches 116; Conservative 58; Mismatches 184; Indels 84; Gaps 18;  
QY 5 SKDAPQADGASGAGQLVPEVNTADPLPMEPVAGPTTAV-----ATAG-----Q 48  
Db 1770 ARTAPQGEAAGTATTASVPGTTTGDG---LDPGVAVTTSVITAENSASIASIATAGIGGPPQQ 1826  
QY 49 VNMIDPWI VNNFVQS PQGEFTTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGVNMR 108  
Db 1827 VDOQETWRTNFYND---VFTWSVADAPGSILYTVQHSFQNNPFTAVLSQMYAGWAGMQ 1883  
QY 109 VRILLAGNAFSAKIIIVCCVPPGF-TSSSLTTIAQATLFPHVADVRLTLEPIEMPLEDVNR 167  
Db 1884 FRFIVAGSVFGRLVRAVIPPGEIIGPGLVRQ---FPHVVIDARSLPEVITITMPDLRP 1940  
QY 168 VLYH-TNDNQPTMRLVCMLYTLRTGGSGNSDSFVAVAGRLTAPSSDFSFLVLP--- 223



```
Db 1941 NMYHPTGDPGLVTLVSVNNLINPFGGSTS---AIQVTVETRPSEDFEFVIRAPSSK 1997
QY 224 TIEQKTRAPTVNPIPLQTLN-SRPSLLIQMLSPDASQVVOFQNGRCLIDQLLGTTP 282
Db 1998 TVDSISPAGLLTPVLTGVGNDNRWNGQIVGLQVPVGGPSTC---NRHNLNGSTVGWSS 2054
QY 283 ATSGQLFRVGKIN-QGARTLNLETV-DGKPFMAFDSP-----APVGFDPFGKCDHMRIS 336
Db 2055 PRFADIDHRGSASYPGSNATVLOFWANAGSAVDNPIISQVAPDGFDP-----MS 2105
QY 337 KTPNNNTSSGDMRVSQVQTVNQVPHLGSIQDFVFNHPTGDIYGTIEWISQSPPTPCT 396
Db 2106 FVPFNGPG-----TPAGWVGFAL-----WNSNGAPNVT 2136
QY 397 DINLWEIPDYGSSLSQANLAP 418
Db 2137 TVQAVEL---GPATGAPGNLQ 2155

RESULT 9
C48562
coat protein - San Miguel sea lion virus (serotype 4)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: C48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of San Miguel sea
eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: C48562
A:Molecule type: genomic RNA
A:Residues: 1-703 <NEI>
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAA16220.1; PID:g334888
A:Note: sequence extracted from NCBI backbone (NCBIP:113567)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.3%; Score 299; DB 1; Length 703;
Best Local Similarity 27.2%; Pred. No. 2.9e-14;
Matches 108; Conservative 60; Mismatches 173; Indels 56; Gaps 17;

QY 4 ASKAPQSDAGSAGQLVPE-----VNTADPLPMEPVAGPTTAVATAGQVNMIDPWVNN 59
Db 151 AESDGP-----GSAEIVTEEQGTVVOQPAP-APTALATLATATGK-SVEQEWMT-- 199
QY 60 FVQSQSGEFTTSPNNTPGDILDLQGLPHNPLFLSHLSQMYNGWGVNMRVILLAGNAPS 119
Db 200 -PFSYHTSINWSTVBSQKILYSQALNPSINPYLDHIAKLYSTWGGIDVRFVTSQGVF 258
QY 120 AGKIIVCCVPBGFTS-SSLTIAQATLFPHVIAVDVTLPIEMPLEDVRNVLVHTNDNQPT 178
Db 259 GGLKALILVPPGVPIESVSMQLQ---YPHVLFDARQTEPVITPIPIRKTLYHSMDDTDTLRLVIVYNEL 315
QY 179 MRLVCLVLTPLRTGSGNSDFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIP 238
Db 316 TKLV---INPYENGVENKTSITVE---TFPSADTFALLKPPGSLIKHGSIFSDILP 368
QY 239 LQTL---SNSRPSLLIQMLSPDASQVVOFQNGRCL-IDGQLLQ-TTPATSGQLFRVGK 294
Db 369 RNSAHMNGNRWMTSTISGFSVQPRV-----FQSNRHFDFDSTTGWSTPYVPVPIEKIQK 423
QY 295 INQARTLNLETVDGKPFMAFDSPAPVGFDPFGKCDHMRISKTNNNTSSGDMRVSQV 354
Db 424 VGSNNKWFHVIDTD-----KALVPGIGDGMPTTIPD---ETKATNGNFSYGESYAGST- 475
QY 355 TNVQGVPHLGSIQDFVFNHPTGDIY-GTIEWISQ 390
Db 476 -----TIKENENSTHFKGYICGTLSVTEIP 501
```

```
RESULT 10
A48562
coat protein - San Miguel sea lion virus (serotype 1)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: A48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of San Miguel sea
eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: A48562
A:Molecule type: genomic RNA
A:Residues: 1-702 <NEI>
A:Cross-references: GB:M87481; NID:g334882; PIDN:AAA16217.1; PID:g334884
A:Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBIP:113565)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.1%; Score 293.5; DB 1; Length 702;
Best Local Similarity 28.3%; Pred. No. 7.4e-14;
Matches 102; Conservative 45; Mismatches 178; Indels 35; Gaps 12;

QY 12 ADGASGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMID-PWIVNMFVQSGEFTI 70
Db 153 SDPGGADIVTEEQGTVVOQPAPQASALTTLAAASTGKTVDCEWTT---PFSYHTAVNM 209
QY 71 SPNTPGDIILDLQGLPHNPLFLSHLSQMYNGWGVNMRVILLAGNAPSAGKIIVCCVPP 130
Db 210 STTEAQQKILFSRALSPELNPLYRHSLSYSGGIDVRFVTSQGVFGVGGKLAALIVPP 269
QY 131 GFTS-SSLTIAQATLFPHVIAVDVTLPIEMPLEDVRNVLVHTNDNQPTMRLVCLMXTPL 189
Db 270 GIEPVESPTMLQ---YPHVLFDARQTEPVITPIPIRKTLYHSMDDTDTLRLVIVYNEL 326
QY 190 RTGGSGNSDFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPLOT--LSNSRF 247
Db 327 INPYEQSBPKS-SCSITVETRPSSDFTFSLKPPGSLIKHGSIPSDILPRNSRHWNGNRW 385
QY 248 PSLIQMLSPDASQVVOFQNGRCL-IDGQLLGTTPATSGQLFRVGKINQARTLNUTE 306
Db 386 WSTIDGFFVQPRV-----FQSNRHFDFDSTTGWSTPYPIPIEVLTLEKDRGGQVFKVTD 440
QY 307 VDGRPFMAPDPAPVGFDPFGKCDHMRISKTNNNTSGD-----PMRSVSQVQNVQGF 360
Db 441 TE-----KSLVP-GLPD-----GWPDITPTAMTASNGNDYTVAEYRITNNGTHFKGF 488

RESULT 11
VCWVFC
coat protein - feline calicivirus (strain Japanese F4)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C:Accession: B40481
R:Toyba, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamaza
Virology 183, 810-814, 1991
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.
A:Reference number: A40481; MUID:91306470; PMID:1853578
A:Accession: B40481
A:Molecule type: genomic RNA
A:Residues: 1-668 <TOH>
A:Cross-references: GB:D90357; NID:g221264; PIDN:BAAL4371.1; PID:g221266
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.8%; Score 284; DB 1; Length 668;
Best Local Similarity 24.7%; Pred. No. 3.5e-13;
Matches 145; Conservative 74; Mismatches 246; Indels 122; Gaps 27;
```

QY 23 PEVNTA-DPLPMPVAGPTTAVATAGVNMIDPWVNNFVQSPQGEFTI-----SPNNT 75  
DB 123 PEQTAAGVGVIAPSAQMSAAMASGKSDVSEW-----EAFSFTSVNWNSTSET 183  
QY 76 PGDILFDLQGLHLPFLSHLSQMYNGWGNMVRILLAGNAFSAAGKIIVCCVPPGFTS- 134  
DB 184 QKILFLKQSLGLPLLPFLHLSKLYVAMSGSIEVRSISGSGVFGKLAIVVPVGDVP 243  
QY 135 SSLTIAQATFLPHVIADVRLTEPIEMPLEDVNRVLYHTNDNQPTMRLVCMLYTPL-RTGG 193  
DB 244 QSTSMQ---YHVLFDARQVPEVFTIIPDLASTLYHVMSTDTTSLVIMVYNDLINPYA 300  
QY 194 GSGNSDSFVAGRVLTAPSSDFELVPPTIEQKTRAFVNPILQTLNS-----R 246  
DB 301 NDSNSGGCIVT--VETKPGDPFKHLLKPP-----GSVLTHGSIIPSDLLPKSSSIWGNR 353  
QY 247 FPLSLQGMILSPDASQVQFQNGRCLIDQQLLGTTPATSGQLFRVRGKINOGARTLNLT 306  
DB 354 YWTDITDFVIRP-----FVFOANRHFDFNQ---ETAGWSTPRFR-----PITITISE 397  
QY 307 VDGKPF---MAFD---SPAPVGFPPDFGKCDWHMRISKTNNNTSSGDPMRS-----V 351  
DB 398 KNGSKLGIGVATDIYIIPGPDGWPDTTIAADKLIPAGDYSITTCGENDIKTAQAYDTAAV 457  
QY 352 SVQTNVQGVPHLGSIQDFEVNHPGTD-YIGTIEWISOP-----STPPGTDINLWEIPD 405  
DB 458 KNTNFRGMV-ICGSLQ-----RAWGDKKISNTAFITTAIRGDNKIEKPSNTIDMTKLAV 510  
QY 406 Y-GSSLSQAANLAPPVFP-----PGFGEAL-----VYFVSAPPFPNNRSA-----PND 447  
DB 511 YQDTHVEQEVQTSDDTLALLGYTGIGBEALGSRDRVRVRSVLPEAGARGGNHPIFYKNS 570  
QY 448 VPCLLPQEVITHFVVSQAQPTMGDAALLHYVDP-----DTNRNLGEFKLYPGGYLT 497  
DB 571 IKLGYVIRSIDVNSQILHTSRQSLNHYLLPDPSPAVYRIIDNSGWSFDDIGIDSEGF-- 628  
QY 498 CVPNGVGA---GQOQLPLNGVFLFVSWVSRYQLKPVGTASTARSRL 541  
DB 629 ---SFVGVSDIGKLEFPLSASYMGI-----QLAKIRLASNIRSRM 665

## RESULT 12

JQ2354  
capsid protein - feline calicivirus (strain NADC)  
C:Species: feline calicivirus  
C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
C:Accession: JQ2354  
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.  
J. Gen. Virol. 74, 2519-2524, 1993  
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable  
A:Reference number: JQ2354; MUID:94065683; PMID:7504075

A:Molecule type: mRNA  
A:Residues: 1-668 <SEA>  
A:Cross-references: GB:L09718; NID:G305104; PIDN:AAA16485.1; PID:G305105  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein

Query Match 9.5%; Score 275.5; DB 2; Length 668;  
Best Local Similarity 25.1%; Pred. No. 1.5e-12;  
Matches 145; Conservative 70; Mismatches 236; Indels 127; Gaps 28;

QY 34 EPVAGPTTAVATAGVNMIDPWVNNFVQSPQGEFTI-----SPNNTPGDILFDLQGLP 87  
DB 145 EPSAQMSAAMATGKSDVSEW-----EAFSFTSVNWNSTSETQKILFKQSLGP 195  
QY 88 HLPFLSHLSQMYNGWGNMVRILLAGNAFSAAGKIIVCCVPPGFTS-SSLTIAQATLPP 146  
DB 196 LLNPLYEHLKLYVAMSGSIEVRSISGSGVFGKLAIVVPVGDVPVQSTSMQ---YP 252  
QY 147 HVIADVRLTEPIEMPLEDVNRVLYHTNDNQPTMRLVCMLYTPL-RTGGSGNSDSFVAG 205  
DB 253 HVLFDARQVPEVIFSIPDLASTLYHLMPTDTTSLVIMVYNDLINPYANDNSNGGCIVT- 311

QY 206 RVLTPAPSSDFSLFVPPPTIEQKTRAFVNPILQTLNS-----RFPSLIQGMILSP 258  
DB 312 -VETKPGDPFKHLLKPP-----GSMITHGVSFSDLLPKSSSLWIGNRYWSDITDFVVRP 365  
QY 259 DASQVQFQNGRCLIDQQLLGTTPATSGQLFRVRGKINOGARTLNLTVEVDGKPF---MAF 315  
DB 366 -----FVFOANRHFDFNQ---ETAGWSTPRFR-----PITITISEKSKLGIGVAT 409  
QY 316 DSPAPVGFPPDFGKCDWHMRISKT-----NNTSSGDPMRSVSV---QTNVQGVFV 361  
DB 410 DYIYIP-GIPD-GWPDTTIAEDLTTPAGYAITSGNGNDITTCSEYDSTDEVIKNNTNFRGM 467  
QY 362 PHLGSIQDFEVNHPGTD-----YIGTI-----EWISQSPSTPPGTDINLWEIPDYGSSL 410  
DB 468 -ICGSLQ-----RAWGDKKISNTAFITTAKEGNKIRPSNTIDMTKIAVYQDTHVGEEV 520  
QY 411 SQAANLAPPVPPPGFGEAL-----VYFVSAPP--GPNNRSAP-----NDVPCLLPQEV 457  
DB 521 QTSDDALLGYTGIGEQAIGSRDRVRVRSVLPEVGARGGNHPIFYKNSIKLGYVIRSI 580  
QY 458 THFVSEQAPTMGDAALLHY-VDPDNRNLGEFKLY-----PGGYLTCCVPNGVG----- 504  
DB 581 DVFNSQILHTSRQSLNHYLLSPDS-----FAVRIIDNSGWSFDDIGIDSEGFSGV 634  
QY 505 -AGPOQLPLNGVFLFVSWVSRYQLKPVGTASTARSRL 541  
DB 635 SIGKLEYPLSASYMGI-----QLAKIRLASNIRSTM 665

## RESULT 13

JQ2356  
capsid protein - feline calicivirus (strain KCD)  
C:Species: feline calicivirus  
C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
C:Accession: JQ2356  
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.  
J. Gen. Virol. 74, 2519-2524, 1993  
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable  
A:Reference number: JQ2354; MUID:94065683; PMID:7504075

A:Molecule type: mRNA  
A:Residues: 1-668 <SEA>  
A:Cross-references: GB:L09719; NID:G305107; PIDN:AAA16487.1; PID:G305108  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein

Query Match 9.5%; Score 275.5; DB 2; Length 668;  
Best Local Similarity 26.0%; Pred. No. 1.5e-12;  
Matches 125; Conservative 58; Mismatches 167; Indels 131; Gaps 24;

QY 34 EPVAGPTTAVATAGVNMIDPWVNNFVQSPQGEFTI-----SPNNTPGDILFDLQGLP 87  
DB 145 EPNAQMSAAMATGKSDVSEW-----EAFSFTSVNWNSTSETQKILFKQSLGP 195  
QY 88 HLPFLSHLSQMYNGWGNMVRILLAGNAFSAAGKIIVCCVPPGFTS-SSLTIAQATLPP 146  
DB 196 LLNPLYEHLKLYVAMSGSIEVRSISGSGVFGKLAIVVPVGDVPVQSTSMQ---YP 252  
QY 147 HVIADVRLTEPIEMPLEDVNRVLYHTNDNQPTMRLVCMLYTPL-RTGGSGNSDSFVAG 205  
DB 253 HVLFDARQVPEVIFSIPDLASTLYHLMSTDTTSLVIMVYNDLINPYANDNSNGGCIVT- 311  
QY 206 RVLTPAPSSDFSLFVLP-----TIEQKTRAFVNPILQTLNSRFPSSLIQGM 254  
DB 312 -VETKPGSDFRPHLLKPPGCVSLTHGVSFSDLLPKTSSSLWIGN-----RYWNITNF 361  
QY 255 ILSPASQVQFQNGRCLIDQQLLGTTPATSGQLFRVRGKINOGARTLNLTVEVDGKPF--- 312  
DB 362 VIRP-----FVFOANRHFDFNQ---ETAGWSTPRFR-----PITITISEQGGTKLGI 405  
QY 313 -MAFSDPAPVGFPPDFGKCDWHMRISKTNNNTSSGDPMRSVSVQTNVQGVPHLGSIQFOE 371

Db 406 GVATDYIVP-GIPD-----GW-----PDTTIS-----EE 428  
QY 372 VFNHPTGDIYGTIEWISQSTPTPGTD-----INLWEIPDYGSSLSQAANLAPVFPFP 423  
Db 429 LI--PAGDYAITNDIGNDITTPAGYDAADTKNNTNFRGMVCG-SLQRA-----475  
QY 424 GGEALV---YFVSAPFGNNNSAPNDVCLLPQEVITHF---VSEQAPTWGDA-ALLH 475  
Db 476 -WGDKKISNTAFITTATVKNKKLPSN---MIDQTKIAVFQDNHVGDKVQTSDDTLALIG 531  
QY 476 Y 476  
Db 532 Y 532

## RESULT 14

VCMWFF  
coat protein - feline calicivirus (strain F9)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C:Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text\_change 28-Jul-2000  
C:Accession: B43382; C45538; PQ0407; S23702  
R:Carver, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.;  
Virology 190, 443-448, 1992  
A:Title: The complete nucleotide sequence of a feline calicivirus.  
A:Reference number: A43382; MUID:92410623; PMID:1529544  
A:Accession: B43382  
A:Molecule type: Genomic RNA  
A:Residues: 1-671 <CARI>  
A:Cross-references: GB:M86379; NID:G323877; PIDN:AAA79327.1; PID:G323879  
R:Carver, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.  
Arch. Virol. 122, 223-235, 1992  
A:Title: Identification and sequence determination of the capsid protein gene of feline  
A:Reference number: A45538; MUID:92117861; PMID:1731695  
A:Accession: C45538  
A:Molecule type: Genomic RNA; protein  
A:Residues: 1-671 <CAR>  
A:Cross-references: GB:M86379; NID:G323877; PIDN:AAA79327.1; PID:G323879  
A:Experimental source: strain F9  
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)  
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.  
J. Gen. Virol. 73, 2429-2433, 1992  
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel  
A:Reference number: PQ0407; MUID:93019069; PMID:1402818  
A:Accession: PQ0407  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>  
A:Cross-references: PID:AB23553.1; PID:G257083  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,304,439,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.3%; Score 268.5; DB 1; Length 671;  
Best Local Similarity 26.8%; Pred. No. 5.1e-12;  
Matches 110; Conservative 47; Mismatches 149; Indels 105; Gaps 18;  
QY 10 QSADGASGAGQLVPEVNT-ADPLPEPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEF 68  
Db 124 EADDSGITA-----PQGTWVGVIAPESQMGTAADMATGKSVDSW-----EAPF 170  
QY 69 TI-----SPNNTPGDILFDLQGLHNPFLSHLSQMYNGWGNMVRILLAGNAPSAGK 122  
Db 171 SPHTSVNWSSTETQKILFKQSLGPLLNPYLEHLAKLYVAMSSIEVRFSGSGVFGCK 230  
QY 123 IIVCCVPPGFTS-SSLTIAQATLPHVIAVDTLSPLEMPLEEDVRNLYHTNDONPTMRL 181  
Db 231 LAAIVVPPGVDPVQSTMLQ---YHVLFDARQVEVPFICLPDLRSLTYHLMSDITDITSL 287  
QY 182 VCMLYTPL-RTGGSGNSDSFVVGRLTAPSSDSFSLFLVPP-----TIEOKT 229  
Db 288 VIMVYNDLINPYANDANSGCCIVT--VETKPGDPDFKHLKPPGSMLTGHSIPSLIPK 345

QY 230 RAFTVNPITLQTLNSRFPFSLQGMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLF 289  
Db 346 SSLWIGN-----RYWSDITDFVIRP-----FVQANRRHDFNQ-----ETAGWSTPRF 308  
QY 290 RYRGKINGQARTLNLTEVDGKFP---MAFDSPAPVGFDPDFGKCDMHMRISKTPNNTSSGD 346  
Db 389 R-----PISVTITEQNGAKLGGVATDYIVP-GIPD-----GW-----PDTTIPGE 428  
QY 347 PMRSVSVQTNVQGVFVPHLGSIOFDEVFNHPTGDIYGTIEWISQSTPTPGTD 397  
Db 429 LI-----PAGDYAITNGTGNDITATGYD 452

## RESULT 15

VCMWFF  
coat protein - feline calicivirus (strain CFI/68 FIV)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C:Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: A40507; B40507; T09246  
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.  
J. Virol. 65, 5440-5447, 1991  
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calici  
A:Reference number: A40507; MUID:91374597; PMID:1716692  
A:Accession: A40507  
A:Molecule type: Genomic RNA  
A:Residues: 1-668 <NEI>  
A:Cross-references: GB:M32819; NID:G323874; PIDN:AAA42925.1; PID:G323875  
A:Accession: B40507  
A:Molecule type: protein  
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>  
R:Neill, J.D.  
A:Title: submitted to the EMBL Data Library, April 1998  
A:Reference number: Z16626  
A:Accession: T09246  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: Genomic RNA  
A:Residues: 1-668 <NE3>  
A:Cross-references: EMBL:U13992; NID:G3056875; PIDN:AAC13993.1; PID:G537256  
A:Experimental source: strain CFI/68 FIV  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.2%; Score 267; DB 1; Length 668;  
Best Local Similarity 31.7%; Pred. No. 6.5e-12;  
Matches 76; Conservative 33; Mismatches 97; Indels 34; Gaps 8;  
QY 34 EPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----SPNNTPGDILFDLQGLP 87  
Db 145 EPNAQMSAADMATGKSVDSW-----EAPSPHTSVNWSSTETQKILFKQSLP 195  
QY 88 HLNPLSHLSQMYNGWGNMVRILLAGNAPSAGKIIIVCCVPPGFTS-SSLTIAQATLPP 146  
Db 196 LLNPYLTLAKLYVAMSGSDVRFSGSGVFGKAAIIVVPPGIDPVQSTMLQ---YP 252  
QY 147 HVIAVDRTLEPIEMPLEDVRNLYHTNDONPTMRLVCMLYTPL-RTGGSGNSDSFVVGAG 205  
Db 253 HVLFDPARQVEVPFISIPDLRSLTYHLMSDITDITSLVIMVYNDLINPYANDANSGCCIVT- 311  
QY 206 RVLTAAPSSDSFSLFVLPPTIEOKTRAFTVNPITLQTLNS-----RPPSLIQGMILSP 258  
Db 312 -VETKPGDPDFKHLKPP-----GSMLTGHSIPSLIPKSSSLWIGNRFSWSDITDFVIRP 365

Search completed: June 1, 2004, 13:55:20  
Job time : 12.5352 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.4195 Seconds  
(without alignments)  
4985.230 Million cell updates/sec

Title: US-09-926-799-7  
Perfect score: 2845  
Sequence: 1 MKMASNDATPSNDGAAGLVLP.....VNQFYSLAPMTGNGRRRVQ 540

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2339	82.2	542	2 S60616	capsid protein - h
2	1870.5	65.7	539	2 S40111	capsid protein - h
3	1229	43.2	530	2 B37471	capsid protein - N
4	1168.5	41.1	546	2 B37491	major capsid prote
5	328	11.5	2344	2 S55399	genome polyprotein
6	320.5	11.3	2344	1 RRWRH	genome polyprotein
7	318	11.2	702	1 A48562	coat protein - San
8	316	11.1	2344	2 S64740	genome polyprotein
9	313	11.0	576	2 A53982	capsid protein - E
10	292.5	10.3	668	1 VCMWF	coat protein - fel
11	292.5	10.3	671	1 VCMWF9	coat protein - fel
12	283.5	10.0	668	2 JQ2354	capsid protein - f
13	280.5	9.9	668	1 VCMWF	coat protein - fel
14	274	9.6	703	1 C48562	coat protein - San
15	269	9.5	668	2 JQ2356	capsid protein - f
16	138	4.9	2333	1 GNNY2F	genome polyprotein
17	137	4.8	2332	1 GNNYF	genome polyprotein
18	136	4.8	1011	1 GNNYC1	genome polyprotein
19	131.5	4.6	733	2 JQ1891	capsid protein - f
20	131	4.6	733	2 JQ1892	capsid protein - f
21	129	4.5	2336	2 S37077	genome polyprotein
22	128.5	4.5	2207	2 S09553	genome polyprotein
23	128	4.5	3624	2 AD0835	large repetitive p
24	127.5	4.5	416	2 T50279	hypothetical serin
25	127	4.5	2332	1 GNNY4F	genome polyprotein
26	126.5	4.4	1978	2 S77257	hypothetical prote
27	124.5	4.4	2205	1 GNNY2W	genome polyprotein
28	123	4.3	2194	1 GNNYF7	genome polyprotein
29	121.5	4.3	3085	2 T00327	polyprotein - infe

30	120.5	4.2	2207	1 GNNY5P	genome polyprotein
31	119	4.2	757	2 JS0198	genome polyprotein
32	118.5	4.2	2206	1 GNNY27	genome polyprotein
33	118	4.1	2209	1 GNNY2P	genome polyprotein
34	117.5	4.1	5376	2 T42215	zonadhesin - mouse
35	115.5	4.1	2175	1 GNNYBE	genome polyprotein
36	114.5	4.0	2206	2 S03822	genome polyprotein
37	114.5	4.0	6658	2 T13931	projectin - fruit
38	114	4.0	2207	1 GNNY1P	genome polyprotein
39	114	4.0	4199	2 S76412	hypothetical prote
40	113	4.0	2185	1 GNNYSV	genome polyprotein
41	112.5	4.0	925	2 S50490	hypothetical prote
42	112.5	4.0	2206	1 GNNY4P	genome polyprotein
43	112	3.9	2209	1 GNNY3P	genome polyprotein
44	109	3.8	257	2 AH0859	periplasmic fibril
45	108.5	3.8	940	2 A40985	projectin - fruit

ALIGNMENTS

RESULT 1

S60616  
capsid protein - human calicivirus (strain Melksham)  
C:Species: human calicivirus  
A:Variety: strain Melksham  
C:Date: 23-May-1997 #sequence\_revision 23-May-1997 #text\_change 28-Jul-2000  
C:Accession: S60616  
R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
Virus Res. 37, 271-283, 1995  
A:Title: capsid glycoprotein in small, round, structured viruses: molecular characterization  
A:Reference numbers: S60616, S60615, S60614, S60613, S60612, S60611, S60610, S60609, S60608, S60607, S60606, S60605, S60604, S60603, S60602, S60601, S60600, S60599, S60598, S60597, S60596, S60595, S60594, S60593, S60592, S60591, S60590, S60589, S60588, S60587, S60586, S60585, S60584, S60583, S60582, S60581, S60580, S60579, S60578, S60577, S60576, S60575, S60574, S60573, S60572, S60571, S60570, S60569, S60568, S60567, S60566, S60565, S60564, S60563, S60562, S60561, S60560, S60559, S60558, S60557, S60556, S60555, S60554, S60553, S60552, S60551, S60550, S60549, S60548, S60547, S60546, S60545, S60544, S60543, S60542, S60541, S60540, S60539, S60538, S60537, S60536, S60535, S60534, S60533, S60532, S60531, S60530, S60529, S60528, S60527, S60526, S60525, S60524, S60523, S60522, S60521, S60520, S60519, S60518, S60517, S60516, S60515, S60514, S60513, S60512, S60511, S60510, S60509, S60508, S60507, S60506, S60505, S60504, S60503, S60502, S60501, S60500, S60499, S60498, S60497, S60496, S60495, S60494, S60493, S60492, S60491, S60490, S60489, S60488, S60487, S60486, S60485, S60484, S60483, S60482, S60481, S60480, S60479, S60478, S60477, S60476, S60475, S60474, S60473, S60472, S60471, S60470, S60469, S60468, S60467, S60466, S60465, S60464, S60463, S60462, S60461, S60460, S60459, S60458, S60457, S60456, S60455, S60454, S60453, S60452, S60451, S60450, S60449, S60448, S60447, S60446, S60445, S60444, S60443, S60442, S60441, S60440, S60439, S60438, S60437, S60436, S60435, S60434, S60433, S60432, S60431, S60430, S60429, S60428, S60427, S60426, S60425, S60424, S60423, S60422, S60421, S60420, S60419, S60418, S60417, S60416, S60415, S60414, S60413, S60412, S60411, S60410, S60409, S60408, S60407, S60406, S60405, S60404, S60403, S60402, S60401, S60400, S60399, S60398, S60397, S60396, S60395, S60394, S60393, S60392, S60391, S60390, S60389, S60388, S60387, S60386, S60385, S60384, S60383, S60382, S60381, S60380, S60379, S60378, S60377, S60376, S60375, S60374, S60373, S60372, S60371, S60370, S60369, S60368, S60367, S60366, S60365, S60364, S60363, S60362, S60361, S60360, S60359, S60358, S60357, S60356, S60355, S60354, S60353, S60352, S60351, S60350, S60349, S60348, S60347, S60346, S60345, S60344, S60343, S60342, S60341, S60340, S60339, S60338, S60337, S60336, S60335, S60334, S60333, S60332, S60331, S60330, S60329, S60328, S60327, S60326, S60325, S60324, S60323, S60322, S60321, S60320, S60319, S60318, S60317, S60316, S60315, S60314, S60313, S60312, S60311, S60310, S60309, S60308, S60307, S60306, S60305, S60304, S60303, S60302, S60301, S60300, S60299, S60298, S60297, S60296, S60295, S60294, S60293, S60292, S60291, S60290, S60289, S60288, S60287, S60286, S60285, S60284, S60283, S60282, S60281, S60280, S60279, S60278, S60277, S60276, S60275, S60274, S60273, S60272, S60271, S60270, S60269, S60268, S60267, S60266, S60265, S60264, S60263, S60262, S60261, S60260, S60259, S60258, S60257, S60256, S60255, S60254, S60253, S60252, S60251, S60250, S60249, S60248, S60247, S60246, S60245, S60244, S60243, S60242, S60241, S60240, S60239, S60238, S60237, S60236, S60235, S60234, S60233, S60232, S60231, S60230, S60229, S60228, S60227, S60226, S60225, S60224, S60223, S60222, S60221, S60220, S60219, S60218, S60217, S60216, S60215, S60214, S60213, S60212, S60211, S60210, S60209, S60208, S60207, S60206, S60205, S60204, S60203, S60202, S60201, S60200, S60199, S60198, S60197, S60196, S60195, S60194, S60193, S60192, S60191, S60190, S60189, S60188, S60187, S60186, S60185, S60184, S60183, S60182, S60181, S60180, S60179, S60178, S60177, S60176, S60175, S60174, S60173, S60172, S60171, S60170, S60169, S60168, S60167, S60166, S60165, S60164, S60163, S60162, S60161, S60160, S60159, S60158, S60157, S60156, S60155, S60154, S60153, S60152, S60151, S60150, S60149, S60148, S60147, S60146, S60145, S60144, S60143, S60142, S60141, S60140, S60139, S60138, S60137, S60136, S60135, S60134, S60133, S60132, S60131, S60130, S60129, S60128, S60127, S60126, S60125, S60124, S60123, S60122, S60121, S60120, S60119, S60118, S60117, S60116, S60115, S60114, S60113, S60112, S60111, S60110, S60109, S60108, S60107, S60106, S60105, S60104, S60103, S60102, S60101, S60100, S60099, S60098, S60097, S60096, S60095, S60094, S60093, S60092, S60091, S60090, S60089, S60088, S60087, S60086, S60085, S60084, S60083, S60082, S60081, S60080, S60079, S60078, S60077, S60076, S60075, S60074, S60073, S60072, S60071, S60070, S60069, S60068, S60067, S60066, S60065, S60064, S60063, S60062, S60061, S60060, S60059, S60058, S60057, S60056, S60055, S60054, S60053, S60052, S60051, S60050, S60049, S60048, S60047, S60046, S60045, S60044, S60043, S60042, S60041, S60040, S60039, S60038, S60037, S60036, S60035, S60034, S60033, S60032, S60031, S60030, S60029, S60028, S60027, S60026, S60025, S60024, S60023, S60022, S60021, S60020, S60019, S60018, S60017, S60016, S60015, S60014, S60013, S60012, S60011, S60010, S60009, S60008, S60007, S60006, S60005, S60004, S60003, S60002, S60001, S60000, S59999, S59998, S59997, S59996, S59995, S59994, S59993, S59992, S59991, S59990, S59989, S59988, S59987, S59986, S59985, S59984, S59983, S59982, S59981, S59980, S59979, S59978, S59977, S59976, S59975, S59974, S59973, S59972, S59971, S59970, S59969, S59968, S59967, S59966, S59965, S59964, S59963, S59962, S59961, S59960, S59959, S59958, S59957, S59956, S59955, S59954, S59953, S59952, S59951, S59950, S59949, S59948, S59947, S59946, S59945, S59944, S59943, S59942, S59941, S59940, S59939, S59938, S59937, S59936, S59935, S59934, S59933, S59932, S59931, S59930, S59929, S59928, S59927, S59926, S59925, S59924, S59923, S59922, S59921, S59920, S59919, S59918, S59917, S59916, S59915, S59914, S59913, S59912, S59911, S59910, S59909, S59908, S59907, S59906, S59905, S59904, S59903, S59902, S59901, S59900, S59899, S59898, S59897, S59896, S59895, S59894, S59893, S59892, S59891, S59890, S59889, S59888, S59887, S59886, S59885, S59884, S59883, S59882, S59881, S59880, S59879, S59878, S59877, S59876, S59875, S59874, S59873, S59872, S59871, S59870, S59869, S59868, S59867, S59866, S59865, S59864, S59863, S59862, S59861, S59860, S59859, S59858, S59857, S59856, S59855, S59854, S59853, S59852, S59851, S59850, S59849, S59848, S59847, S59846, S59845, S59844, S59843, S59842, S59841, S59840, S59839, S59838, S59837, S59836, S59835, S59834, S59833, S59832, S59831, S59830, S59829, S59828, S59827, S59826, S59825, S59824, S59823, S59822, S59821, S59820, S59819, S59818, S59817, S59816, S59815, S59814, S59813, S59812, S59811, S59810, S59809, S59808, S59807, S59806, S59805, S59804, S59803, S59802, S59801, S59800, S59799, S59798, S59797, S59796, S59795, S59794, S59793, S59792, S59791, S59790, S59789, S59788, S59787, S59786, S59785, S59784, S59783, S59782, S59781, S59780, S59779, S59778, S59777, S59776, S59775, S59774, S59773, S59772, S59771, S59770, S59769, S59768, S59767, S59766, S59765, S59764, S59763, S59762, S59761, S59760, S59759, S59758, S59757, S59756, S59755, S59754, S59753, S59752, S59751, S59750, S59749, S59748, S59747, S59746, S59745, S59744, S59743, S59742, S59741, S59740, S59739, S59738, S59737, S59736, S59735, S59734, S59733, S59732, S59731, S59730, S59729, S59728, S59727, S59726, S59725, S59724, S59723, S59722, S59721, S59720, S59719, S59718, S59717, S59716, S59715, S59714, S59713, S59712, S59711, S59710, S59709, S59708, S59707, S59706, S59705, S59704, S59703, S59702, S59701, S59700, S59699, S59698, S59697, S59696, S59695, S59694, S59693, S59692, S59691, S59690, S59689, S59688, S59687, S59686, S59685, S59684, S59683, S59682, S59681, S59680, S59679, S59678, S59677, S59676, S59675, S59674, S59673, S59672, S59671, S59670, S59669, S59668, S59667, S59666, S59665, S59664, S59663, S59662, S59661, S59660, S59659, S59658, S59657, S59656, S59655, S59654, S59653, S59652, S59651, S59650, S59649, S59648, S59647, S59646, S59645, S59644, S59643, S59642, S59641, S59640, S59639, S59638, S59637, S59636, S59635, S59634, S59633, S59632, S59631, S59630, S59629, S59628, S59627, S59626, S59625, S59624, S59623, S59622, S59621, S59620, S59619, S59618, S59617, S59616, S59615, S59614, S59613, S59612, S59611, S59610, S59609, S59608, S59607, S59606, S59605, S59604, S59603, S59602, S59601, S59600, S59599, S59598, S59597, S59596, S59595, S59594, S59593, S59592, S59591, S59590, S59589, S59588, S59587, S59586, S59585, S59584, S59583, S59582, S59581, S59580, S59579, S59578, S59577, S59576, S59575, S59574, S59573, S59572, S59571, S59570, S59569, S59568, S59567, S59566, S59565, S59564, S59563, S59562, S59561, S59560, S59559, S59558, S59557, S59556, S59555, S59554, S59553, S59552, S59551, S59550, S59549, S59548, S59547, S59546, S59545, S59544, S59543, S59542, S59541, S59540, S59539, S59538, S59537, S59536, S59535, S59534, S59533, S59532, S59531, S59530, S59529, S59528, S59527, S59526, S59525, S59524, S59523, S59522, S59521, S59520, S59519, S59518, S59517, S59516, S59515, S59514, S59513, S59512, S59511, S59510, S59509, S59508, S59507, S59506, S59505, S59504, S59503, S59502, S59501, S59500, S59499, S59498, S59497, S59496, S59495, S59494, S59493, S59492, S59491, S59490, S59489, S59488, S59487, S59486, S59485, S59484, S59483, S59482, S59481, S59480, S59479, S59478, S59477, S59476, S59475, S59474, S59473, S59472, S59471, S59470, S59469, S59468, S59467, S59466, S59465, S59464, S59463, S59462, S59461, S59460, S59459, S59458, S59457, S59456, S59455, S59454, S59453, S59452, S59451, S59450, S59449, S59448, S59447, S59446, S59445, S59444, S59443, S59442, S59441, S59440, S59439, S59438, S59437, S59436, S59435, S59434, S59433, S59432, S59431, S59430, S59429, S59428, S59427, S59426, S59425, S59424, S59423, S59422, S59421, S59420, S59419, S59418, S59417, S59416, S59415, S59414, S59413, S59412, S59411, S59410, S59409, S59408, S59407, S59406, S59405, S59404, S59403, S59402, S59401, S59400, S59399, S59398, S59397, S59396, S59395, S59394, S59393, S59392, S59391, S59390, S59389, S59

[illegible]

RESULT 2

S40111 capsid protein - human calicivirus (strain Bristol isolate B493)  
C:Species: human calicivirus  
A:Variety: strain Bristol isolate B493  
C:Date: 25-Dec-1994 #sequence\_revision 27-Feb-1997 #text\_change 28-Jul-2000  
C:Accession: S40111  
R:Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
A:Description: The EMBL Data Library, December 1993  
A:Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de  
A:Reference number: S40111  
A:Accession: S40111  
A:Molecule type: genomic RNA  
A:Residues: 1-539 <GRE>  
A:Cross-references: EMBL:X76716; NID:g436410; PIDN:CAA54134.1; PID:g436411  
A:Experimental source: Human enteric calicivirus strain Bristol isolate B493  
C:Superfamily: human calicivirus capsid protein  
C:Keywords: capsid protein; coat protein

Query Match	65.7%	Score 1870.5	DB 2	Length 539
Best Local Similarity	64.3%	Pred. No. 1.2e-129		
Matches 348	Conservative 82	Mismatches 104	Indels 7	Gaps 4
QY	1	MKMASNDATPSNDGAAGLVPSNNEAMALEPVVGASLAAPVTGTQNIIDPWIRTNFVQAP	60	
Db	1	MKMASNDANPDSGGSANLVPEVNNEVMALEPVVGAIAAPVAGQONVIDPWIRNFEVQAP	60	
QY	61	NGEFTVSPRNPGEILVNLGPELNPYLAHARMYNGYAGGMEVQVWLAGNAFTAGKII	120	
Db	61	GGEFTVSPRNAPGETLWGAFLGPDLPNLPYLSHLSRMYNGYAGGFEVQVTLAGNAFTAGKVI	120	
QY	121	FAAVPPYPPEVNLSPSQITMPHPHVIDRTLEPVLPMDDVRSRTLHFNQKDEPKORLVA	180	
Db	121	FAAVPPPNPFTGLSPSQVTMPFHIIIVDRVQLEPVLIPLDVRNRYFHYNQANDSTLKIIA	180	
QY	181	MLYTPLRNGSGDDVFTVSCRILTRPSPEFDFTYLVPTTVESKTKPFTLPVLTLGELSNS	240	
Db	181	MLYTPLRNAGDDVFTVSCRVLTRPSDFDFILVPPTVESRTKPTFPVPLTVBEMSNS	240	
QY	241	RPPLSIDEMVTPSNESI VVQPONGRVTLDELLGTTQLQACNICSIRGKVTQGVPSBOHM	300	
Db	241	RPPIPLEKLYTGFSSAFVVPQNGRCRTTDGVLGTTQLSAVNICNFRGDDVTHIAGS--HD	298	
QY	301	WNLEITNLNGTQFDTDVDPALPGVDPDFAGEVFGVLSQRNRCESNPANRAHDVAVTATYSD	360	
Db	299	YTMNLASQWSNYDPTSEELPAPLGPDPFGVKIQLGLTQTTRADG--STFAHKATYSTGSV	356	
QY	361	KYTPKLGVLQIGTWTNNDVE-NQPTKFTPIGLNEVANGHRFE--QWTLPRYSGALTUNNN	417	
Db	357	HTFPKLSGVQFTTDNNDFQAGNTKFTPVGVIQGDHHDHNEPQQWLLPNYSGRGTGHNVH	416	
QY	418	LAPAVAPLPGBRLLLPFRSYVPLKGGFGNPAIDCSVPQEWVHOHFQESAPSLGDVALVRY	477	
Db	417	LAPAVAPTFPGQLLFFFRSTMFGCSGYPNMNLDCLLPQEWVLHFYQEAAPASQDVALLRF	476	
QY	478	VNPDTGRVLFEAKLHKGGFLTVSSSTGTGPVWPVANGYFKFDSWVNOFYSLAPMGTCNGRR	537	
Db	477	VNPDTGRVLFECKLHKSGIITVAHTGPDYDLVLPNGYFRFDSWVNOFYTLAPMGNTGRR	536	
QY	538	R	538	

Db 537 R 537

RESULT 3

B37471  
capsid protein - Norwalk virus  
C/Species: Norwalk virus  
C/Date: 24-Feb-1994 #sequence\_revision 25-Apr-1997 #text\_change 28-Jul-2000  
C/Accession: B37471  
R/Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.  
Virology 195, 51-61, 1993  
A/Title: Sequence and genomic organization of Norwalk virus.  
A/Reference number: A37471; MUID:93303939; PMID:8391187  
A/Accession: B37471  
A/Status: preliminary; not compared with conceptual translation  
A/Molecule type: genomic RNA  
A/Residues: 1-530 <JIA>  
A/Cross-references: GB:M87661; NID:g1061311; PIDN:AAB50466.1; PID:g1061313  
A/Note: sequence extracted from NCBI backbone (NCBIP:134157)  
C:Superfamily: human calicivirus capsid protein

Query Match	43.2%	Score 1229	DB 2	Length 530
Best Local Similarity	47.4%	Pred. No. 1.6e-82		
Matches 262	Conservative	82	Mismatches 163	Indels 46
Gaps				
QY	1	MMKASNDATPSNDCAAG---	LVPESN-NEAMALEPVPVVGASLAAPVTCGTNIIDPIWRTNF	56
DB	1	MMMASKDATSSVDCASGAGQLVPEVNADPLAMPVAGSSATAVATAGQVNPIDPWIINNF	60	
QY	57	VOAPNGFTVSPRNSPGEIILVNLELGPENAPYLAAHARMYNGVYAGGMEVQVMLAGNAFTA	116	
DB	61	VOAPQGGFTTSPNNIPGVDVLPDLGLGPHLNPFLHLQMYNGWVGNMRVIRLMLAGNAFTA	120	
QY	117	GKIIIFAAVPPYPVENLSPSGITMPPHVIDVRLTEPVLPLMPDVRSTLPHFNQKDEPKM	176	
DB	121	GKIIIVSGIPPGFGSHLTTIAQATTFPHVIAADVRLTDPIEVELEDVRNVLFENNDRNQOTM	180	
QY	177	RLVAMLYTPLRNSGDDVFTVSCRIILTRPSPBDFTVLPVPTVESKTKPTFLPVLTLGE	236	
DB	181	RLVCMLYTPLRGTGGTDSFVVVAGRVMTCPSPDFENFLVLPVPTVEQTRPTFLPNLPLSS	240	
QY	237	LSNGRFPPLSIDEMWTPSNESIVQOPNGRVTLDBELLGTTQLQACNICSIRGKVTQVPS	296	
DB	241	LSNSRAPLPISSIGISPDNVQSVQFQNGRCITDGLRVLGTTVPVLSHVAKIRGTSNGTV--	298	
QY	297	EQHMNLNLTINLNGTQFDPTDDVPAPLVDPDFAG-----EVFGVLISQRNGESNPANR	349	
DB	299	-----INLTLDGTPFPHPFEG-PAPIGFDPLGGCDWHINMTQFHSQTO-----	342	
QY	350	AHDAAVATYSKYTPKGLVQ---IGTWNTNDVENQPTKFTPIGLNEVANGHRFPQWTLPL	406	
DB	343	-YD--VDTPPTDFVPHLGSIQANGISGNGYGV-----LSWISPPSHSPSQVDLWKIP	393	
QY	407	RYSGALTLMNMLAPAVAPLFPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWVQHYPQBSA	466	
DB	394	NYGSSITEATHLAPSVYPPGFEVLVFFMSKMPGPGAYNLP---CLLPQEYISHLASBOA	450	
QY	467	PSLGDVALVRYNPDGTGVLPEAKLHKGGLFIV--SSTSTGVPVVVPANGYKFXDGSWNQF	524	
DB	451	PTVGAALLUHYVDPDTGRNLGFEKAYPDGFLTCTCVNGASSGQQPLPINGVFVFWVSWSRF	510	
QY	525	YSLAPMGTGNRR	537	
DB	511	YOLKPVGTASSAR	523	

RESULT 4  
B37491  
major capsid protein [similarity] - Southampton virus  
N;Alternate names: orf2 protein  
C;Species: Southampton virus  
C;Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #t

A;Residues: 1-2344 <ROS>  
A;Cross-references: EMBL:X87607; NID:G854640; PIDN:CAA60910.1; PID:G854641  
A;Experimental source: isolate BS89  
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C;Keywords: polyprotein

Query Match 11.5%; Score 328; DB 2; Length 2344;  
Best Local Similarity 26.1%; Pred. No. 2.6e-15;  
Matches 155; Conservative 77; Mismatches 210; Indels 152; Gaps 30;  
Qy 3 MASNDATPSNDGAGL-----VPESNNEAMALEPVV-----GASLAAPVTG-- 43  
Db 1766 MEKARTAPQGEAAGATTASTVPGITTD--GLDPGVVATTSVTAENSSASTAIGGP 1923  
Qy 44 --QTNIIDPIWRTNFVQAPNGBFTVSPRNSPGEILVNLELGPENLPYLAAHARMYNGYAG 101  
Db 1824 PQVDQOETW-RTNFY--NDVFTWSVAADAPGSILTVQHSQNNPFATVLSQMYAGWAG 1980  
Qy 102 GMEVQVNLGAGTAFTAKIIPAAVPPVPVENLSPS-QITMFPHVIIDVLTLEPLVLPMPD 160  
Db 1881 GMPREITVAGSGVFGRLVAIVPGI--EIGPGLEVRQFPHVVIDARSLSEPVITITMPD 1937  
Qy 161 VRSTLPHFQKQBPKM--RLVAMLYTPLRASGGDDVFTVSCRILTRPSPEEDFTLVVP 218  
Db 1938 LRPNMYH--PTDGPGLVPTLVLVSVNNL-INPFGSTSAIQVTETRPSDEFVVMIRAP 1994  
Qy 219 ---TVESKTKP--FTLPVLTLGLSLNSRRPLSIDEMVTSPNESIV-VQPNQGRVT----- 267  
Db 1995 SSKTVDSISPGALLITPVLT-GVGNDRW-----NGQIVGLQVPVGGFSTCNRH 2042  
Qy 268 --LDGELLGTTQLQACNICSIRKGVTVQVPSQHMNNLBITNLNGTFQDPTDDVPAPLGV 325  
Db 2043 WNLNGSTYGSSPRFADIDHRRGSASYPGSNATNVLQVFNAGSAVDNPISQV-APDGF 2101  
Qy 326 PDF-----AGEV-FGVLSORNGESNPAN-RAHDVVAT----- 357  
Db 2102 PDMSFVPFNGPGIPAGAGVGFCAIMNSNSGAPNVTTVQAYELGFGATGAPCNLOPTTNSG 2161  
Qy 358 ---YSDKYTPKGLVQ-----IGTWNTNDVENQPT----- 384  
Db 2162 AQTVAKSIYAVVTGTAQNAGLFWMASGVISTPNANAITYTPQDRIVTTPGTTPAAPVG 2321  
Qy 385 KFTPIGLNEV-----ANGHRFEQMTLPYSGALTLNNMLAPAVAPLFFGERL 431  
Db 2222 KNTPIFASVVRTGDVNATAGSANGTYGTSQPL---PVTIGLSLNNYSALMFGQ-- 2276  
Qy 432 LFRFSYVPLKGGFNPAIDCSVPQEWQHFYQESAPS-----IGDVALVRYNVP 480  
Db 2277 -FFVMQLTTFASG--MEIGLSVD---GVFYAGTGASTTLIDLTELIDVRVPGP 2323

RESULT 6  
RRWRRH  
genome polyprotein - rabbit hemorrhagic disease virus  
N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C;Species: rabbit hemorrhagic disease virus  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998  
C;Accession: A41039  
R;Meyers, G.; Wirblich, C.; Thiel, H.J.

```

RRWRWH
genome polypeptin - rabbit hemorrhagic disease virus
N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)
C;Species: rabbit hemorrhagic disease virus
C;Date: 30-Sep-1992 #sequence_revision 30-Sep-1992 #text_change 29-May-1998
C;Accession: A41039
R;Meyers, G.; Wirblich, C.; Thiel, H.J.
Virology 184, 664-676, 1991
A;Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide seq
A;Reference number: A41039; MUID:91361557; PMID:1840711
A;Accession: A41039
A;Molecule type: Genomic RNA
A;Residues: 1-2344 <MEY>
A;Cross-references: GB:M67473
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C;Keywords: nucleotidyltransferase

Query Match      11.3%; Score 320.5; DB 1; Length 2344;
Best Local Similarity 25.9%; Pred. No. 9.3e-15;
Matches 152; Conservative Indels 147; Gaps 30;

```

QY 8 ATPSDNGA---AGLVPSNNEAM-----ALEPVV-----CASLAAPVTG-----QTNIID 49  
 DB 1772 AAPQEAAGATTASVPGTGTGMDGPGVAVATTSVITAENSASATAGIGPPQOVQOE 1831  
 QY 50 PWIRTNFQAPNGEFTVSPRNSPGELLVNLGPELNPYLAHLARMYNGYAGGMEVQVML 109  
 DB 1832 TW-RTNFY--NDVFTWSADAPGILYTVQHSQNNPFTAVLSQYAGNAGGQFPRIV 1888  
 QY 110 AGNAFTAGKIIFAAAPPVPFVBNLSPS-QITMFPHVIIIDVRLTEPVLPMPEVRSSTLPHF 168  
 DB 1889 AGSGVFGKRLVRAVIPPDI---EIGGLEVRQFPHVIDARSLEPVTITPDLRPNMYH- 1944  
 QY 169 NQKDBPKM--RLVAMLYTFLRNSGGDDVFTVSCRILTRPSPEPFTYLVPP---TVESK 223  
 DB 1945 -PTGDDPLVPTLVLSVYNL--INPFGGSTSAIQVTETRPSSDFFVMIRAPSSKTVDSI 2002  
 QY 224 TKP--FTLPLVLITGELNSRFPISIDEMVTSNESIV-VOPQNGRVT-----LDGELL 273  
 DB 2003 SPAGLLTTPLT--GVGNDRW-----NGQIVGLQPVPGFSTCNRHNLNGSTY 2050  
 QY 274 GTTOLQACNICIRGKVTGQVPSEQHMNLEITNLNGTQFDPTDDVPAPLGVPDF----- 328  
 DB 2051 GMSPPFGIDHRRGSASVSGSNATNLQFMYANAGSAIDNFIQV-APDGPDMSFVPF 2109  
 QY 329 -----AGEV-FGVLSQRNRESNPAN-RAHDAVVAT-----YSDK 361  
 DB 2110 NGRGIPAAAGWVGFAIWNNSGAPNVTTVQAVELGATGAPGNLOPTTNTSQAQTVAKSI 2169  
 QY 362 YTPKGLQVQ-----ICTWNTNDVENOPT-----KFTPIGLN 392  
 DB 2170 YAVVTGAQNPAGLFWNASGIIITPNAQRIITYTPQDRIVTTPGTAAAPVGKNTPIMPA 2229  
 QY 393 EV-----ANGHREQWTLPRYSGALTNNMLAPAVAPLFGGERLLFFRSYVP 439  
 DB 2230 SVVVRTGDVNATAGSANGTQGTGSQLP-----PVTIGLSLNNYSALMPGQ-----FFVWQLT 2283  
 QY 440 LKGFGNPAIDCSVPQEWVHFQESAPS-----LGDVALRVYVNP 480  
 DB 2284 FASGF--MEIGLSVD-----GYFAGTGASTLIDLDELIDVRVGP 2323  
 RESULT 7  
 A48562  
 coat protein - San Miguel sea lion virus (serotype 1)  
 N;Alternate names: capsid protein  
 C;Species: San Miguel sea lion virus  
 C;Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
 C;Accession: A48562  
 R;Neill, J.D.  
 Virus Res. 24, 211-222, 1992  
 A;Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lions.  
 A;Reference number: A48562; MUID:92410750; PMID:1529644  
 A;Accession: A48562  
 A;Molecule type: genomic RNA  
 A;Residues: 1-702 <NE>  
 A;Cross-references: GB:M87481; NID:g334882; PIDN:AAA16217.1; PID:g334884  
 A;Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBI:P:113565)  
 C;Superfamily: feline calicivirus coat protein  
 C;Keywords: capsid protein; coat protein; glycoprotein  
 F;208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 Query Match 11.2%; Score 318; DB 1; Length 702;  
 Best Local Similarity 25.8%; Pred. No. 2.4e-15;  
 Matches 153; Conservative 74; Mismatches 199; Indels 168; Gaps 33;  
 QY 8 ATPSDNGAAGLVPSNNEAMALEPVVVGAS-----LAAPVTGQTNIIID-PWIRTNFVQ---A 59  
 DB 151 AESDGGGADIIVTESEGTVQQQPVPAQSALTTLAAAGTGKT--VDCW--TTFFSYHTA 206  
 QY 60 PNGEFTVSPRNSPGELLVNLGPELNPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKI 119  
 DB 207 VNWSTT-----EAQKILFSTRALSPELNPLYRHSLSYTSWGGIDVRFTVSGSVFGKGL 262

QY 120 IFAAAPPVF-PVENLSPSOITMFPHVIIIDVRLTEPVLPMPEVRSSTLPHFNQKDEPKMRL 178  
 DB 263 AALIVPPGEPVE--SPTML-QYPHVLFDARQTEPVIPTIDIRKTLVH-SMDDTDTTRL 318  
 QY 179 VAMLYTFL-RNSGGDDVFTVSCRILTRPSPEPFTYLVPPVESKTKPFTLPLVLTGL 237  
 DB 319 VIMVYNELINPYEQSEPKSSCSITVETRPSSDFTSLKPPG-----SLKHGSI 368  
 QY 238 SNRFPPLSIDEMVTSNESIVVQPNQR-----VTLDGELLGTTQLQACNICIRK 289  
 DB 369 -----PSDLI---PRNSRHMGNRWSTIDGFFVQVPRVQSNRHFDFDST 410  
 QY 290 VTG-----QVPSRQHMNLEITNLNGTQFDPTDD---VPA-PLGVDP----- 327  
 DB 411 TTGWSPTIPIEV---TLEKLDRGQYFKAVTDTEKSLVPEGLPDGWPDTTIPTAMTASNG 467  
 QY 328 -----FAG-EVFGVLSQR-----NRGESNPANRAHDAVVATYSOKY-- 362  
 DB 468 NYDYTVAEYRITNNGTHFKGYIMGNLTTKVKGSDNLGETQOTSRTLFLASVGNKYKQNTI 527  
 QY 363 --TPKGLVVOIGTWNTNDVENOPTKT-----PIGLN-----EVANGH 398  
 DB 528 NPHKITNSLVVDANNVSAATAKTTTWHSTMSHLGYVLVDESPPVGSDDSTKVRIA--- 584  
 QY 399 RFEQWTLPRYSGALTNNMLAPAVAPLFGGERLLFFRSYVPLKGGFGNPAIDC-----S 452  
 DB 585 -----TLPE---AFTNGN-----FP-----VFTNKIQI-GHFDRAHTKCFNSQVLM 623  
 QY 453 VPQEWVHFQESAPSGLDVALRVYVNPDTGRVLFEAKLHKGGLFLTVSSSTSGP 506  
 DB 624 TSQKLAENHY-----TLPPDSLLVIRITDAASSWFDLGINHDFSYV-GISTIP 671  
 RESULT 8  
 S64740  
 genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)  
 N;Contains: VP60 protein  
 C;Species: rabbit hemorrhagic disease virus  
 A;Variety: isolate AST/89  
 C;Date: 12-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 20-Jun-2000  
 C;Accession: S64740; S46944; S49018; S65012  
 R;Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.  
 submitted to the EMBL Data Library, May 1995  
 A;Description: Genomic organization of rabbit hemorrhagic disease virus determined by d.  
 A;Reference number: S64740  
 A;Accession: S64740  
 A;Molecule type: genomic RNA  
 A;Residues: 1-2344 <CAS>  
 A;Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552  
 A;Experimental source: isolate AST/89  
 R;Boga, J.; Casais, R.; Martin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.  
 submitted to the EMBL Data Library, July 1993  
 A;Description: Molecular cloning, sequence and expression of the capsid protein gene fro  
 A;Reference number: S46944  
 A;Accession: S46944  
 A;Molecule type: Genomic RNA  
 A;Residues: 1650-2344 <BOG>  
 A;Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623  
 A;Experimental source: isolate AST/89  
 R;Parra, F.; Boga, J.A.; Martin, M.S.; Casais, R.  
 Virus Res. 27, 219-228, 1993  
 A;Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp  
 A;Reference number: S49018; MUID:93255896; PMID:8488721  
 A;Accession: S49018  
 A;Molecule type: Genomic RNA  
 A;Residues: 1650-1796 <PAR>  
 A;Cross-references: EMBL:Z24757  
 A;Experimental source: isolate AST/89  
 A;Accession: S65012  
 A;Molecule type: protein  
 A;Residues: 1767-1779;1875-1877, 'X',1940-1941 <PAW>  
 C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase



C:Keywords: blocked amino end; polyprotein

Query Match 11.1%; Score 316; DB 2; Length 2344;  
Best Local Similarity 25.7%; Pred. No. 2e-14;  
Matches 152; Conservative 80; Mismatches 212; Indels 148; Gaps 30;

QY 3 MASNDATPSNDCAGL-----VPESNNEAM-----ALEPV-----GASLAAPVTG-----43  
DB 1766 MEGKARTAPQGAAGATTASVPGTTDGMDFGWATTSVTAENSSASIAATAGTGGPPQ 1825

QY 44 QTNIIIDPWIRTNFVQAPNCEFTVSPRNSPGEILVNLGPELNPYLALHARMYGVAGGM 103  
DB 1826 QVDQOETW-RTFYF--NDVFTWSVADAPGSLITYVQHSFQNNPFTAVLSQMYAGWAGM 1882

QY 104 EYQVMLAGNAFTAGKIIFAAPVPPYFVENLSPS-QITMFPFHVIIIDVRLLEPVLMPDVR 162  
DB 1883 QFRFIVAGIVGEGRLVAAPFPGI---BIGPLEVRQFPHVVIDARSLEPVTITMPDLR 1939

QY 163 STLPHFNOKDEPKM--RLVAMLYTPLRSNGSGDDVFTVSCRILTRPSPEFDTYLVPP-- 218  
DB 1940 PNMVH--PTGDPLGVTPLVLSVNNL-INPFGGSTSAIQVTVETRPSEDFEFVIRAPSS 1996

QY 219 -TVESKTKP--FTLPLVLTGELSNRFPPLSIDEMWTSNPSIV-VQPNQGRVT-----267  
DB 1997 KTVDSISPAGLITTPVLT-GVGNDRW-----NGQIVGLQVPVGGSTCNRHN 2044

QY 268 LDGELLGTTQLOACNICSRKGVTVQVPSQEQHMNLEITNLNGTQFDPTDDVPAPLGVDP 327  
DB 2045 LNSTGVWSSPRFGDIGHRRGASVPGNNATVLPWFYANAGSAIDNPISQV-APDGPD 2103

QY 328 F-----AGEV-FGVLSQNRGESN-----PAN-----348  
DB 2104 MSFVPFNGGPIPAAGWVGFAIWNNSGAPNVTTQAYELGFATGAPGNLQPTTNSGQ 2163

QY 349 ---RAHDAVATYSK---YTPKGLVQIGTMTNDVENOPTKE-----TPIGLN 392  
DB 2164 TVAKSIYAVVTQAQNPAGLFWAGSVITSPSANAITYTPQPDRIVTPGTAAAPVGN 2223

QY 393 E-----VANGHRFEQWTLPRYSGALTLMNMLAPAVAPLFGGERLLF 433  
DB 2224 TPIMPASVVRRTGDVNTAGSANGTYGTGSQPL---PVTIGLSLNNYSSALMPQ---P 2277

QY 434 FRSYVPLKGGFNGPAIDCSVPOEWQHFQESAPS-----LGDVALRVYVNP 480  
DB 2278 FVWQLTFASGF--MBEIGLSVD---GYFYAGTGASTLIDLTELIDVVRPVG 2323

RESULT 9  
A:Accession: A53982  
A:Title: capsid protein - European brown hare syndrome virus  
C:Species: European brown hare syndrome virus  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 28-Jul-2000  
C:Accession: A53982  
R:Wiethrich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Eakens, U.; Haas, B.; Thiel, H.  
J. Virol. 68, 5164-5173, 1994  
A:Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease  
A:Reference number: A53982; MUID:94309183; PMID:7518531  
A:Accession: A53982  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-576 <WR>  
A:Cross-references: GB:U09199  
C:Superfamily: human calicivirus capsid protein

Query Match 11.0%; Score 313; DB 2; Length 576;  
Best Local Similarity 26.6%; Pred. No. 4.1e-15;  
Matches 111; Conservative 67; Mismatches 148; Indels 92; Gaps 21;

QY 3 MASNDATPSNDCAGLVPESNNEAMALEPVVGASLAAPVTGNTIIDPWIRTNFVQAPNG 62  
DB 31 VASTDVTADNVAS-----VATAGTGGP-PQASQESW-RVNFYF--ND 72

QY 63 EFTVSPRNSPGEILVNLGPELNPYLALHARMYGVAGMEVQVMLAGNAFTAGKIIFA 122

DB 73 VFTVSWTDAPGSILYSVQHSPPQNNPFTQVLSQMYAGWAGMQFRFIVAGSGIFGGRLVCA 132  
QY 123 AVPPYFPVENLSPS-QITMFPFHVIIIDVRLLEPVLMPDVRSTLFHFNOKDEPKM--RLV 179  
DB 133 ITPPGIQIQ---PGLEVRQFPHVVIDARSLEPVTITMPDLRPEMYH--PTGDFGLVPTLV 187

QY 180 AMLYTPPLRSNGSGDDVFTVSCRILTRPSPEFDTYLVPP---TVESKTKP--TLPVLTL 234  
DB 188 VSVNNL-INPFGGSTSAIQVTVETRPSEDFEVLIRAPSSKTVDSVNSPMLLTTPVLT- 245

QY 235 GELSNRFPPLSIDEMWTSNPSIVQV-----ONGRVTLDGELLGTTQLOACNICSR 287  
DB 246 GAGSDNRWAGPI-----VGLQVPVGGSTSNRHNMGSTYGMSSPRFPDDIDHPS 295

QY 288 GKV---TGQVPSQEQHMNLEITNLNGTQFDPTDDVPAPLGVDP-----PAGEV-----332  
DB 296 GNVSYPTGSATNTIETW---YANAGTATNPISNI-APDGFDMGAIPIPSGGTITPTGAMV 351

QY 333 -FGVLSQNRGESNPNRAHDAVATYSKYPKGLVQIGTMTNDVENOPTKFTPI 389  
DB 352 GFGQVWNASG-----TPYVGTQV---AYELGFANGAGPSIRPV 387

RESULT 10  
VCWVFF  
coat protein - feline calicivirus (strain CFI/68 FIV)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: A40507; B40507; T09246  
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.  
J. Virol. 65, 5440-5447, 1991  
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calicivirus  
A:Reference number: A40507; MUID:91374597; PMID:1716692  
A:Accession: A40507  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <NE1>  
A:Cross-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875  
A:Accession: B40507  
A:Molecule type: protein  
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>  
R:Neill, J.D.  
submitted to the EMBL Data Library, April 1998  
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.  
A:Reference number: Z16626  
A:Accession: T09246  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <NE3>  
A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256  
A:Experimental source: strain CFI/68 FIV  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
P:177,301,304,439,459,615/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 10.3%; Score 292.5; DB 1; Length 668;  
Best Local Similarity 26.0%; Pred. No. 1.6e-13;  
Matches 113; Conservative 53; Mismatches 143; Indels 125; Gaps 21;

QY 11 SNDGAAGLVPPESNN--EAMALEPVVGASLAAPVTGNTIIDPW-IRTNFVQAPNGEFTVS 67  
DB 125 ADDGSI-TTPEGTVMVGVIAPNAQMTAAADMATGKSDVSEWAEFFSFHTSVNN---S 179

QY 68 PRNSPGEILVNLGPELNPYLALHARMYGVAGMEVQVMLAGNAFTAGKIIFAAPVPPY 127  
DB 180 TSETQGTGLFKQSLGPLNPYLTLAKLYVAVSGSDVRFISGSGVFGGKLAALVWPPG 239

QY 128 F-PVENLSPSQTTFPHVVIDVRLLEPVLMPDVRSTLFHFNOKDEPKMRLVAMLYTPL 186  
DB 240 IDPQVSTMLQ---YPHVLFDARQVEPVIFSIDPLRSTLYHL-MSDPTTSLVIMVNDL 295

QY 187 -----RNSGSGDDVFTVSCRILTRPSPEFDTYLVPPVTVESKTKPFTLPLVLTGELS 240

Db 296 INPYANDSSSG-CIVTVE---TKPGDFKPHLLKPPG-----SMLTHGSI--- 337  
 QY 241 REPLSIDEMWTSNPSIIVVQPNQGRVTLTGELLGTTQLQACNICSIIRGKVTGQVPSQOHM 300  
 Db 338 -----PSDLI-----PKSSSL 348  
 QY 301 WNLEITNLNGTQF--DPTDDVPAPLGVDPDFAGEVFGV-----LSQRNGESNPANRAHDA 353  
 Db 349 W-----IGNRFWSIDITFVIRPF-----VFOANRHFDFNQETAGWSTPRFR---P 390  
 QY 354 VVATVSDKYTPKLGIVQIGT-----WNTNDVENQPTKFTPIGLNEVANGHRPEQW 403  
 Db 391 ITITISVKESAKLG-IGVATDYIVFGIPDGMPTTI---PGLVFPVGDVAITNGNNDIT 446  
 QY 404 TLPYSGALTLNWN 417  
 Db 447 TAAQYDAATEIRNN 460

RESULT 11  
 VCWHP9  
 coat protein - feline calicivirus (strain F9)  
 N;Alternate names: capsid protein  
 C;Species: feline calicivirus  
 C;Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 28-Jul-2000  
 C;Accession: B43382; C45538; PQ0407; S23702  
 R;Carter, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.  
 Virology 190, 443-448, 1992  
 A;Title: The complete nucleotide sequence of a feline calicivirus.  
 A;Reference number: A43382; MUID:92410623; PMID:1529544  
 A;Accession: B43382  
 A;Molecule type: genomic RNA  
 A;Residues: 1-671 <CAR1>  
 A;Cross-references: GB:M86379; NID:G323877; PIDN:AAA79327.1; PID:G323879  
 R;Carter, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.  
 Arch. Virol. 122, 223-235, 1992  
 A;Title: Identification and sequence determination of the capsid protein gene of feline  
 A;Reference number: A45538; MUID:92117861; PMID:1731695  
 A;Accession: C45538  
 A;Molecule type: genomic RNA; protein  
 A;Residues: 1-671 <CAR2>  
 A;Cross-references: GB:M86379; NID:G323877; PIDN:AAA79327.1; PID:G323879  
 A;Experimental source: strain F9  
 A;Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)  
 R;Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.  
 J. Gen. Virol. 73, 2429-2433, 1992  
 A;Title: The cloning, sequencing and expression of a major antigenic region from the fel  
 A;Reference number: PQ0407; MUID:93019069; PMID:1402818  
 A;Accession: PQ0407  
 A;Status: nucleic acid sequence not shown  
 A;Molecule type: mRNA  
 A;Residues: 217-266, 'SI', 269-336, 'Y', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>  
 A;Cross-references: PIDN:AA823553.1; PID:G257083  
 C;Superfamily: feline calicivirus coat protein  
 C;Keywords: capsid protein; coat protein; glycoprotein  
 F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.3%; Score 292.5; DB 1; Length 671;  
 Best Local Similarity 22.7%; Pred. No. 1.7e-13;  
 Matches 133; Conservative 84; Mismatches 223; Indels 147; Gaps 25;

QY 11 SNDGAAGLVPSNN--EAWALEPVVGASIAAPVGTQTNIDPW-IRTNFVQAPNGEFTVS 67  
 Db 125 ADDGSI-TAPEQGTWVGVIAPSAQMTAADMATGKSVDSWEAFSFTSVNW----S 179  
 QY 68 PRNSPGEILVNLGPELNPYLALHARMYNGYAGMGVEQVQVLMAGNAFTAGKIIFAAPPPY 127  
 Db 180 TSETQKILFKQSLGPLLNPYLEHLAKLVAMSGSIEVRFSISGSGVFGGKLAIVVPP- 238  
 QY 128 FVENLSPSQITWTFPHVIDVTLRVLVLPMDVSTLTFHFNQKDEPKMRLVAMLYTLR 187  
 Db 239 -GVDPVQSTSMQLQPHVLFDAQVBEVIFCLPDLRSLTYHL--MSDITDTSVLVWYNDL- 295

QY 188 SNGSGDDVFTVSC--RIILTRPSPBFDFTYLVPP-----TVESKTKPFTLPVLTLGELS 238  
 Db 296 INPYANDSSSGCIVTVEKPGDFKPHLLKPPGSMTHGSIPLKPT-SSLWIG--- 351  
 QY 239 NSRFPLSIDEMVTSP-----NE-----SIVVQPQNGR----- 265  
 Db 352 -NRYWSDITDFVIRPPVFOANRHFDFNQETAGWSTPRFRPISVTTTEONGAKLGIGVATD 410  
 QY 266 -----VTLGELLGTTQLQACN-----ICSIIRGKVTGQV-----P 295  
 Db 411 YIVPGIPDGMPTTIPEGLIPAGDVAINGTGNDITATGYDTADIIKNTNFRGMVIG 470  
 QY 296 SEQHMWNLEITNLNGTQFDPDTPDDVPAPLGVDPDFAGEVFGVLSQRNGESNPANRAHDAV 355  
 Db 471 SLORAWGDK--KISNTAFIT-----ATLGDNNKINPCNTIDQSKI 511  
 QY 356 ATYSDKYTPKLGIVQIGTWNNDVENQPTKFTPIGLNEVANGHRPEQ-----WTLPRYSGA 411  
 Db 512 VWFQDNHVGKKA-----QTSDDTLALIGYT--GIGEQAIGSDRDRVVRISTLPE-TGA 561  
 QY 412 LTLNMLAPAVAPLPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWVQHFYQESAPSLGD 471  
 Db 562 R-----GGNHPIFYKNSIKL--GYVIRSIDV-----FNSQILHTRSQ 596  
 QY 472 VALRVYVNPDTGRVLFPEAKLHKGGLTVSTSTGTPVVPVANGYFKFD 518  
 Db 597 LSLNHYLLPPDPSFAVYRIIDNSGWFIDIGIDSDGFSFVGVSGFGKLE 643

RESULT 12  
 JQ2354  
 capsid protein - feline calicivirus (strain NADC)  
 C;Species: feline calicivirus  
 C;Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
 C;Accession: JQ2354  
 R;Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.  
 J. Gen. Virol. 74, 2519-2524, 1993  
 A;Title: Analysis of feline calicivirus capsid protein genes: identification of variable  
 A;Reference number: JQ2354; MUID:94065683; PMID:7504075  
 A;Accession: JQ2354  
 A;Molecule type: mRNA  
 A;Residues: 1-668 <SEA>  
 A;Cross-references: GB:L09718; NID:G305104; PIDN:AAA16485.1; PID:G305105  
 C;Superfamily: feline calicivirus coat protein  
 C;Keywords: capsid protein; coat protein

Query Match 10.0%; Score 283.5; DB 2; Length 668;  
 Best Local Similarity 23.6%; Pred. No. 7.6e-13;  
 Matches 130; Conservative 85; Mismatches 224; Indels 111; Gaps 24;

QY 30 EFVVGASLAAPVGTQTNIDPW-IRTNFVQAPNGEFTVSPRNSPGEILVNLGPELNPY 88  
 Db 145 EFSAQMTAADMATGKSVDSWEAFSFTSVNW---STSETQKILFKQSLGPLLNPY 200  
 QY 89 LAHLARMYNGYAGMGVEQVQVLMAGNAFTAGKIIFAAPPPVFPVENLSPSQITWTFPHVIDV 148  
 Db 201 LEHLSKLYVAMSGSVEVRFSISGSGVFGGKLAIVVPP--GVDPVQSTSMQLQPHVLFDA 258  
 QY 149 RTLEPVLLPMDVRSSTLTFHFNQKDEPKMRLVAMLYTLP-----RSNGSGDDVFTVSCRI 202  
 Db 259 RQVDFVIFSIPLDRSLTYHL-MPDITDTSVLVWYNDLINPYANDSSSG-CIVTVE--- 313  
 QY 203 LTRPSPBFDFTYLVPTVESKTKPFTLPVLTLGELSNRFPPLS-----IDEMVTSNPSIV 258  
 Db 314 -TKPGDFKPHLLKPPG-----SMLTHGSPVSDIIPKSSSLWIGNRWTSIDTPV 362  
 QY 259 VQP-----QNGRVTLDCELLGTTQLQ-----ACNICSIIRGKVTG-----QVPSQOHM-- 301  
 Db 363 VRFVFOANRHFDFNQETAGWSTPRFRPITITISSKGLGIGVATDYIVPGIPDGMWD 422  
 QY 302 -----NLEITNLN-----QTQDPTDDVPAPLGVDPDFAGEVFGVLSQRNGES 344

Db 423 TTIAEDLTGAGDYATGNGNDITTTGSEYDSTEVIK---NNTNFRGMVYICGSLQRAWGDK 479  
QY 345 NPANRAHDAVAVATYDKYTPK--LGLVQIGTWTNDV-ENQTKFTPI-----GLNEVA 395  
Db 480 KISNTAFITTAKEGKIRPSNTIDMTKIAVYQDTHVGEVQTSDDALALLGYTGGEQA 539  
QY 396 NGRHREQWTLPRYSALTNNMLAPAVAPLFFGERLLFRSVVPLKGGFGNPAIDCSVPQ 455  
Db 540 IGSDRDR-----VVRISVLPEVGAR-CGNHPIFYKNSIKL--GVVIRSIDV---- 582  
QY 456 EWWQHFYOBAPSGLGDVALRY-VNPDTRGLV-----FEAKLHKGGLUTVSSTNG 505  
Db 583 -----FNSQILHTSRLQSLNHLNLLSPDFAVVRIIDNSGWFIDIGSEGFVGISSIG 637  
QY 506 PVVVPANGYF 515  
Db 638 KLEYPLSASY 647

RESULT 13  
VCWVFC  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: B40481  
R:Tohya, Y.; Taniguchi, Y.; Takahashi, E.; Utogawa, E.; Takeda, N.; Miyamura, K.; Yamaza  
Virology 183, 810-814, 1991  
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.  
A:Reference number: A40481; MUID:91306470; PMID:1853578  
A:Accession: B40481  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <TOH>  
A:Cross-references: GB:D90357; NID:g221264; PIDN:BAAL4371.1; PID:g221266  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 280.5; DB 1; Length 668;  
Best Local Similarity 32.0%; Pred. No. 1.3e-12;  
Matches 78; Conservative 36; Mismatches 93; Indels 37; Gaps 9;

QY 9 TPSNDGAGLVPESSNNEAMALEPVVGSAAAPVQTQNTIIDPW-IRTNFVQAPNGEFTVS 67  
Db 132 TPEQGTAVG-----GVIAEPQAQMSTADMASGKSVDSWEAFPSFHTSVNW----S 179  
QY 68 PRNSPGEILVNLELGPENLPYLAHLARMYNGYAGMEVQVMLAGNAFTAGKIIFAAVPPY 127  
Db 180 TSETQKILFKQSLGRLNPLNLYLHSLKLYVAMSGSIEVRFSGSGVFGGKLAALVWPP- 238  
QY 128 FPVENLSPSQTMTFPHVVIDVRLPMPDVRSTLPHFNQKDEPKMRLVAMLYTPI- 186  
Db 239 -GVDPVQSTMLQYPHVLFARQVPEVFTIPDLKSTLYHV-MSDITDTSLSVIMYNDLI 296  
QY 187 -----RNSGSGDDVFTVSCRILTRPSPEDFDFTLVLPVPTVESKTPFTLPLVLTGLSLSNR 241  
Db 297 NPVANDSNSG-CIVTVE---TKGPDQFKPHLLKPPG-----SVLTHGSIPLSDL 341  
QY 242 FPLS 245  
Db 342 IPKS 345

RESULT 14  
C48562  
N:Alternate names: capsid protein  
C:Species: San Miguel sea lion virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
C:Accession: C48562  
R:Neill, J.D.  
Virus Res. 24, 211-222, 1992  
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel

eins.  
A:Reference number: A48562; MUID:92410750; PMID:1529644  
A:Accession: C48562  
A:Molecule type: genomic RNA  
A:Residues: 1-703 <NEI>  
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888  
A>Note: sequence extracted from NCBI backbone (NCBIP:113567)  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.6%; Score 274; DB 1; Length 703;  
Best Local Similarity 28.4%; Pred. No. 4.1e-12;  
Matches 93; Conservative 54; Mismatches 142; Indels 38; Gaps 14;

QY 8 ATPSNDGAGLVPESSNNEAMALEPVVGSAAAPVQTQNTIIDPW-IRTNFVQAPNGEFTVS 63  
Db 151 AESDGPSSAEIVTEQGTAVVQQAAPAPATALATATATGK-SVEQEMW-TFF--SYHTS 206  
QY 64 FTVSPRNSPGEILVNLELGPENLPYLAHLARMYNGYAGMEVQVMLAGNAFTAGKIIFAA 123  
Db 207 INWSTVESQKILYVAMSGSIEVRFSGSGVFGGKLAALVWPP- 266  
QY 124 VPP-YPPVENLSPSQTMTFPHVVIDVRLPMPDVRSTLPHFNQKDEPKMRLVAMLY 182  
Db 267 VPPGVEPIESVSMQLQ---YPHVLFARQVPEVFTIPDLKSTLYHV-MSDITDTSLSVIM 322  
QY 183 YPLRSNGSGDDVFTVSCRILTRPSPEDFDFTLVLPVPTVESKTPFTLPLVLTGLSLSNR 240  
Db 323 Y-----ENGVTCSITVETRPSSADFTFALLKPPGSLIKHGSIPSDILPRNSAHMGN 377  
QY 241 RPLSIDEMVTSPNSISIVVQPNQGRVTLTGELLGTTQLQACNICIRGKVTQGVPSBOHM 300  
Db 378 RWSSTISGFSVQPR---VFSNRHFDST---TTGWSPTPVYVPIEIKIQGKVSNNKW 430  
QY 301 WNLEITNLNGTQFDPTDDVPAPLGPVD 327  
Db 431 FHV-----IDTDKALVP-GIPD 446

RESULT 15  
JQ2356  
capsid protein - feline calicivirus (strain KCD)  
C:Species: feline calicivirus  
C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
C:Accession: JQ2356  
R:Seal, B.S.; Ridpath, J.P.; Mengeling, W.L.  
J. Gen. Virol. 74, 2519-2524, 1993  
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable  
A:Reference number: JQ2354; MUID:94065683; PMID:7504075  
A:Accession: JQ2356  
A:Molecule type: mRNA  
A:Residues: 1-668 <SEA>  
A:Cross-references: GB:L09719; NID:g305107; PIDN:AAAL6487.1; PID:g305108  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein

Query Match 9.5%; Score 269; DB 2; Length 668;  
Best Local Similarity 23.9%; Pred. No. 8.8e-12;  
Matches 137; Conservative 83; Mismatches 234; Indels 120; Gaps 28;

QY 11 SNDGAGLVPESSN--EAMALEPVVGSAAAPVQTQNTIIDPW-IRTNFVQAPNGEFTVS 67  
Db 125 ADDGSI-VTPEQGTAVGVIAPNAQMSTADMATKSVDSWEAFPSFHTSVNW----S 179  
QY 68 PRNSPGEILVNLELGPENLPYLAHLARMYNGYAGMEVQVMLAGNAFTAGKIIFAAVPPY 127  
Db 180 TSETQKILFKQSLGRLNPLNLYLHSLKLYVAMSGSIEVRFSGSGVFGGKLAALVWPP- 238  
QY 128 FPVENLSPSQTMTFPHVVIDVRLPMPDVRSTLPHFNQKDEPKMRLVAMLYTPI- 186  
Db 239 -GVDPVQSTMLQYPHVLFARQVPEVFTIPDLKSTLYHL-MSDITDTSLSVIMYNDLI 296

Search completed: June 1, 2004, 13:55:27  
 . Job time : 12.4195 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 31.4396 Seconds  
(without alignments)  
5419.275 Million cell updates/sec

Title: US-09-926-799-7  
Perfect score: 2845  
Sequence: 1 MKNASNDTPSNDGAAGLVP.....VNOFYSLAPMTGNGRRRVQ 540

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.25.\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mnc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_rvirus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2811	98.8	540	Q917X1	Q917x1 norwalk-lik
2	2806	98.6	538	Q91I65	Q91i65 human calic
3	2796	98.3	540	Q91I7W8	Q91i7w8 norwalk-lik
4	2789	98.0	540	Q91V50	Q91v50 human calic
5	2361	83.0	542	Q12305	Q12305 snow mounta
6	2361	83.0	542	Q80RD6	Q80rd6 snow mounta
7	2360	83.0	542	Q96877	Q96877 snow mounta
8	2351	82.6	542	Q913B6	Q913b6 human calic
9	2339	82.2	542	Q66293	Q66293 calicivirus
10	2326	81.8	548	Q915C7	Q915c7 human calic
11	2273.5	79.9	535	Q91I15	Q91i15 human calic
12	2270.5	79.8	535	Q68104	Q68104 hawaii cali
13	2269.5	79.8	535	Q9QMK6	Q9qmk6 chitta viru
14	2268.5	79.7	535	Q915C6	Q915c6 human calic
15	2268.5	79.7	535	Q917X7	Q917x7 norwalk-lik
16	2261.5	79.5	535	Q917X4	Q917x4 norwalk-lik

17	2261.5	79.5	535	12	Q916E4	Q916e4 human calic
18	2260.5	79.5	535	12	Q8V784	Q8v784 norwalk vir
19	2258.5	79.4	535	12	Q917Y3	Q917y3 norwalk-lik
20	2257.5	79.3	535	12	Q916E6	Q916e6 human calic
21	2254.5	79.2	535	12	Q917Y9	Q917y9 norwalk-lik
22	2250.5	79.1	535	12	Q917Y0	Q917y0 norwalk-lik
23	2238.5	78.7	535	12	Q91V51	Q91v51 human calic
24	2219.5	78.0	535	12	Q91V39	Q91v39 human calic
25	2026.5	71.2	547	12	Q8V773	Q8v773 norwalk vir
26	2023	71.1	548	12	Q8V0P5	Q8v0p5 human calic
27	2023	71.1	548	12	Q88291	Q88291 small round
28	2023	71.1	548	12	Q8V0P4	Q8v0p4 human calic
29	2021	71.0	548	12	Q917Z2	Q917z2 norwalk-lik
30	2021	71.0	548	12	Q66296	Q66296 calicivirus
31	2021	71.0	548	12	Q8V768	Q8v768 norwalk vir
32	2020	71.0	548	12	Q8BC87	Q8bc87 human calic
33	2018.5	70.9	547	12	Q8V775	Q8v775 norwalk vir
34	2018	70.9	540	12	Q918B0	Q918b0 norwalk-lik
35	2015	70.8	548	12	Q916E5	Q916e5 human calic
36	2014	70.8	548	12	Q8BC90	Q8bc90 human calic
37	2014	70.8	548	12	Q80FL1	Q80fl1 human calic
38	2014	70.8	548	12	Q80FK4	Q80fk4 human calic
39	2014	70.8	550	12	Q918B3	Q918b3 norwalk-lik
40	2013	70.8	548	12	Q918A1	Q918a1 norwalk-lik
41	2012.5	70.7	547	12	Q918A7	Q918a7 norwalk-lik
42	2010	70.7	548	12	Q9PYA7	Q9pya7 human calic
43	2010	70.7	548	12	Q917Z5	Q917z5 norwalk-lik
44	2010	70.7	548	12	Q80FK9	Q80fk9 human calic
45	2010	70.7	548	12	Q80FK8	Q80fk8 human calic

ALIGNMENTS

RESULT 1

Q917X1 ID Q917X1 PRELIMINARY; PRT; 540 AA.  
AC Q917X1;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Capsid protein.  
GN ORF2.  
OS Norwalk-like virus NLV/New Orleans/306/1994/US.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=171854;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/New Orleans/306/1994/US;  
RX MEDLINE=97193806; PubMed=9041391;  
RA Ando T., Montoe S.S., Noel J.S., Glass R.I.;  
RT A one-tube method of reverse transcription-PCR to efficiently amplify  
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
RL of small round-structured viruses (Norwalk-like viruses).";  
RL ~~GenBank:U00001.153570-577(1992)~~  
RN [2]  
RP SEQUENCE OF 100-192 FROM N.A.  
RC STRAIN=Hu/NLV/New Orleans/306/1994/US;  
RX MEDLINE=98071177; PubMed=9407386;  
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,  
RA Seto Y., Montoe S.S., Glass R.I.;  
RT A correlation-of-patient immune responses with genetically  
RT characterized small round-structured viruses involved in outbreaks of  
RT nonbacterial acute gastroenteritis in the United States, 1990 to  
RT 1995. ~~Med Virol 65:372-385(1992)~~  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/New Orleans/306/1994/US;  
RX MEDLINE=20266071; PubMed=10804147;  
RA Ando T., Noel J.S., Fankhauser R.L.;  
RT "Genetic classification of 'Norwalk-like viruses.'";

J. Infect. Dis. 181:S336-S348 (2000) .  
 [4]  
 SEQUENCE FROM N.A.  
 STRAIN=Hu/NLV/New Orleans/306/1994/US;  
 Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;  
 Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.  
 EMBL; AF414422; AAL13019.1; -  
 InterPro; IPR004005; Calici\_coat.  
 InterPro; IPR008975; Viral\_Cap\_coat.  
 Pfam; PF00915; Calici\_coat\_1.  
 SEQUENCE 540 AA; 5866 MW; F0D0B52BBC0B6024 CRC64;  
 Query Match 98.8%; Score 2811; DB 12; Length 540;  
 Best Local Similarity 98.9%; Pred. No. 3e-214;  
 Matches 534; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 QY 1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVGASLAAPVTGQTNIIDPWIRTFVQAP 60  
 DB 1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVGASLAAPVTGQTNIIDPWIRTFVQAP 60  
 QY 61 NGEFTVSPRNSPGEILVNLELGPENLYLAHLARMYNGYAGGMEVQVNLGNAFTAGKII 120  
 DB 61 NGEFTVSPRNSPGEILVNLELGPENLYLAHLARMYNGYAGGMEVQVNLGNAFTAGKII 120  
 QY 121 FAAVPPYPVENLSPSQITMFPFHVIDVRLTLEPVLPMDDVRSITLPHFNKDEPKMLVA 180  
 DB 121 FAAVPPYPVENLSPSQITMFPFHVIDVRLTLEPVLPMDDVRSITLPHFNKDEPKMLVA 180  
 QY 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDFTYLPVPTVESKTKPFTLPVLTGLSLSNS 240  
 DB 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDFTYLPVPTVESKTKPFTLPVLTGLSLSNS 240  
 QY 241 RPLSIDEMVTSPNESIVVQPNQGRVTLTGELGTTQLOACNICSIKRGKVTGVPSQHM 300  
 DB 241 RPLSIDEMVTSPNESIVVQPNQGRVTLTGELGTTQLOACNICSIKRGKVTGVPSQHM 300  
 QY 301 WNLEITNLNGTQFDPTDDVPALGVPDFAGEVFGVLSQRNGESNPANRAHADVAATYSD 360  
 DB 301 WNLEITNLNGTQFDPTDDVPALGVPDFAGEVFGVLSQRNGESNPANRAHADVAATYSD 360  
 QY 361 KYTPKLGVLQIGTWNINDVENQPTKFTPIGLNEVANGHREFQWTLPRYSGALTINMNLAP 420  
 DB 361 KYTPKLGVLQIGTWNINDVENQPTKFTPIGLNEVANGHREFQWTLPRYSGALTINMNLAP 420  
 QY 421 AVAPLPPGERLLAFRRSYVPLKGGFGNPAIDCSVPQEWQHFQYESAPSLGDVALVRYNP 480  
 DB 421 AVAPLPPGERLLAFRRSYVPLKGGFGNPAIDCLVPQEWQHFQYESAPSLGDVALVRYNP 480  
 QY 481 DTGRVLFEAKLHKGGLTVSSTSTGPPVVVPANGYKFDSSWNQFYSLAPMGTGNGRRVQ 540  
 DB 481 DTGRVLFEAKLHKGGLTVSSTSTGPPVVVPANGYKFDSSWNQFYSLAPMGTGNGRRVQ 540  
 RESULT 2  
 Q91165 PRELIMINARY; PRT; 538 AA.  
 ID AC Q91165  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Capsid protein.  
 OS Human calicivirus NLV/MOH/99.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OC NCBI\_TaxID=168224;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HU/NLV/MOH/99;  
 RC Farkas T., Berke T., Reuter G., Szucs G., Matson D., Jiang X.;  
 RT "Molecular detection and sequence analysis of human caliciviruses from  
 RT acute gastroenteritis outbreaks in Hungary."  
 RT Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RA Farkas T., Berke T., Reuter G., Szucs G., Matson D., Jiang X.,  
 RA NCBI\_TaxID=168224;  
 RA [1]  
 RA SEQUENCE FROM N.A.  
 RA STRAIN=HU/NLV/MOH/99;  
 RA Farkas T., Berke T., Reuter G., Szucs G., Matson D., Jiang X.;  
 RA "Molecular detection and sequence analysis of human caliciviruses from  
 RA acute gastroenteritis outbreaks in Hungary."  
 RA Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 RA NCBI\_TaxID=168224;  
 RA [1]  
 RA SEQUENCE FROM N.A.  
 RA STRAIN=HU/NLV/White River/290/1994/US;  
 RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Bates M.K.,  
 RA NCBI\_TaxID=171855;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=97193806; PubMed=9041391;  
 RX Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
 RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
 RT of small round-structured viruses (Norwalk-like viruses).";  
 RL J. Clin. Microbiol. 35:570-577 (1997).  
 [2]  
 RN SEQUENCE OF 100-192 FROM N.A.  
 RP STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=98071277; PubMed=9407386;  
 RX Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Bates M.K.,  
 RA NCBI\_TaxID=171855;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=97193806; PubMed=9041391;  
 RX Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
 RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
 RT of small round-structured viruses (Norwalk-like viruses).";  
 RL J. Clin. Microbiol. 35:570-577 (1997).  
 [2]  
 RN SEQUENCE OF 100-192 FROM N.A.  
 RP STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=98071277; PubMed=9407386;  
 RX Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Bates M.K.,  
 RA NCBI\_TaxID=171855;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=97193806; PubMed=9041391;  
 RX Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
 RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
 RT of small round-structured viruses (Norwalk-like viruses).";  
 RL J. Clin. Microbiol. 35:570-577 (1997).  
 [2]  
 RN SEQUENCE OF 100-192 FROM N.A.  
 RP STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=98071277; PubMed=9407386;  
 RX Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Bates M.K.,  
 RA NCBI\_TaxID=171855;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=97193806; PubMed=9041391;  
 RX Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
 RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
 RT of small round-structured viruses (Norwalk-like viruses).";  
 RL J. Clin. Microbiol. 35:570-577 (1997).  
 [2]  
 RN SEQUENCE OF 100-192 FROM N.A.  
 RP STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=98071277; PubMed=9407386;  
 RX Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Bates M.K.,  
 RA NCBI\_TaxID=171855;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=97193806; PubMed=9041391;  
 RX Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
 RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
 RT of small round-structured viruses (Norwalk-like viruses).";  
 RL J. Clin. Microbiol. 35:570-577 (1997).  
 [2]  
 RN SEQUENCE OF 100-192 FROM N.A.  
 RP STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=98071277; PubMed=9407386;  
 RX Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Bates M.K.,  
 RA NCBI\_TaxID=171855;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=97193806; PubMed=9041391;  
 RX Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
 RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
 RT of small round-structured viruses (Norwalk-like viruses).";  
 RL J. Clin. Microbiol. 35:570-577 (1997).  
 [2]  
 RN SEQUENCE OF 100-192 FROM N.A.  
 RP STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=98071277; PubMed=9407386;  
 RX Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Bates M.K.,  
 RA NCBI\_TaxID=171855;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=97193806; PubMed=9041391;  
 RX Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
 RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
 RT of small round-structured viruses (Norwalk-like viruses).";  
 RL J. Clin. Microbiol. 35:570-577 (1997).  
 [2]  
 RN SEQUENCE OF 100-192 FROM N.A.  
 RP STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=98071277; PubMed=9407386;  
 RX Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Bates M.K.,  
 RA NCBI\_TaxID=171855;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/White River/290/1994/US;  
 RC MEDLINE=97193806; PubMed=9041

RA Seto Y., Monroe S.S., Glass R.I.;  
 RT "Correlation of patient immune responses with genetically  
 RT characterized small round-structured viruses involved in outbreaks of  
 RT nonbacterial acute gastroenteritis in the United States, 1990 to  
 RT 1995."; J. Med. Virol. 53:372-383(1997).  
 RL J. Med. Virol. 53:372-383(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/White River/290/1994/US;  
 RX MEDLINE=20266071; PubMed=10804147;  
 RA Ando T., Noel J.S., Fankhauser R.L.;  
 RT "Genetic classification of 'Norwalk'-like viruses."; J.  
 RL J. Infect. Dis. 181:5336-5348(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/White River/290/1994/US;  
 RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF414423; AAL13022.1; -;  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 540 AA; 58788 MW; 9C8164BEA52F53B9 CRC64;

Query Match 98.3%; Score 2796; DB 12; Length 540;  
 Best Local Similarity 98.5%; Pred. No. 4.7e-213;  
 Matches 532; Conservative 1; Mismatches 7; Indels 0; Gaps 0;  
 QY 1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVGASLAAPVTGQTNIDPWIRTFVQAP 60  
 DB 1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVGASLAAPVTGQTNIDPWIRTFVQAP 60  
 QY 61 NGEFTVSPNSPGEILVNLLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKII 120  
 DB 61 NGEFTVSPNSPGEILVNLLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKII 120  
 QY 121 FAAPPPYFVENLSPSQITMFPFHVIDVRLPEVLLPMPDVRSSTLPHFNQKDEPKMLVA 180  
 DB 121 FAAPPPYFVENLSPSQITMFPFHVIDVRLPEVLLPMPDVRSSTLPHFNQKDEPKMLVA 180  
 QY 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEFDTYLPVPTVESKTKPFTLPVLTGLGELNS 240  
 DB 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEFDTYLPVPTVESKTKPFTLPVLTGLGELNS 240  
 QY 241 RPLSLIDEMVTSFNESIVVQPNQGRVTLTGELGTTQLQACNICSIKRGVTCQVPSQHM 300  
 DB 241 RPLSLIDEMVTSFNESIVVQPNQGRVTLTGELGTTQLQACNICSIKRGVTCQVPSQHM 300  
 QY 301 WNLEITNLNGTQFDDTDDVPAPLGVDPDFAGEVGVLSQRNRESNPANRAHDAVVATYSD 360  
 DB 301 WNLEITNLNGTQFDDTDDVPAPLGVDPDFAGEVGVLSQRNRESNPANRAHDAVVATYSD 360  
 QY 361 KYTPKLGIVQIGTWNNDVENOPTKFTPIGLNEVANGHRFEQWTLPRYSGALTLNNLAP 420  
 DB 361 KYTPKLGIVQIGTWNNDVENOPTKFTPIGLNEVANGHRFEQWTLPRYSGALTLNNLAP 420  
 QY 421 AVAPLPPGERLLFFRSYVPLKGGFNPALDCSVPOEWQHFYQESAPSLGDVALVRYNVP 480  
 DB 421 AVAPLPPGERLLFFRSYVPLKGGFNPALDCSVPOEWQHFYQESAPSLGDVALVRYNVP 480  
 QY 481 DTGRVLFEAKLHKGFLTVSSTSTGPPVVVPANGYKFDSSWNQFYS LAPMGTCNGRRRVQ 540  
 DB 481 DTGRVLFEAKLHKGFLTVSSTSTGPPVVVPANGYKFDSSWNQFYS LAPMGTCNGRRRVQ 540

RESULT 4  
 Q9IV50  
 ID Q9IV50 PRELIMINARY; PRT; 540 AA.  
 AC Q9IV50;  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Capsid protein.

OS Human calicivirus HU/NLV/Hillingdon/90/UK.  
 OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OX NCBI\_TaxID=122911;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/Hillingdon/90/UK;  
 RX MEDLINE=20404883; PubMed=10949950;  
 RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,  
 RA Clegg J.C., Chamberlain J., Brown D.W.G.;  
 RT "Capsid protein diversity among 'Norwalk'-like viruses."; J.  
 RL Virus Genes 20:227-236(2000).  
 DR EMBL; AJ277607; CAB89088.1; -;  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 540 AA; 59043 MW; B224257E68B0DDCD CRC64;

Query Match 98.0%; Score 2789; DB 12; Length 540;  
 Best Local Similarity 98.1%; Pred. No. 1.7e-212;  
 Matches 530; Conservative 4; Mismatches 6; Indels 0; Gaps 0;  
 QY 1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVGASLAAPVTGQTNIDPWIRTFVQAP 60  
 DB 1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVGASLAAPVTGQTNIDPWIRTFVQAP 60  
 QY 61 NGEFTVSPNSPGEILVNLLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKII 120  
 DB 61 NGEFTVSPNSPGEILVNLLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKII 120  
 QY 121 FAAPPPYFVENLSPSQITMFPFHVIDVRLPEVLLPMPDVRSSTLPHFNQKDEPKMLVA 180  
 DB 121 FAAPPPYFVENLSPSQITMFPFHVIDVRLPEVLLPMPDVRSSTLPHFNQKDEPKMLVA 180  
 QY 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEFDTYLPVPTVESKTKPFTLPVLTGLGELNS 240  
 DB 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEFDTYLPVPTVESKTKPFTLPVLTGLGELNS 240  
 QY 241 RPLSLIDEMVTSFNESIVVQPNQGRVTLTGELGTTQLQACNICSIKRGVTCQVPSQHM 300  
 DB 241 RPLSLIDEMVTSFNESIVVQPNQGRVTLTGELGTTQLQACNICSIKRGVTCQVPSQHM 300  
 QY 301 WNLEITNLNGTQFDDTDDVPAPLGVDPDFAGEVGVLSQRNRESNPANRAHDAVVATYSD 360  
 DB 301 WNLEITNLNGTQFDDTDDVPAPLGVDPDFAGEVGVLSQRNRESNPANRAHDAVVATYSD 360  
 QY 361 KYTPKLGIVQIGTWNNDVENOPTKFTPIGLNEVANGHRFEQWTLPRYSGALTLNNLAP 420  
 DB 361 KYTPKLGIVQIGTWNNDVENOPTKFTPIGLNEVANGHRFEQWTLPRYSGALTLNNLAP 420  
 QY 421 AVAPLPPGERLLFFRSYVPLKGGFNPALDCSVPOEWQHFYQESAPSLGDVALVRYNVP 480  
 DB 421 AVAPLPPGERLLFFRSYVPLKGGFNPALDCSVPOEWQHFYQESAPSLGDVALVRYNVP 480  
 QY 481 DTGRVLFEAKLHKGFLTVSSTSTGPPVVVPANGYKFDSSWNQFYS LAPMGTCNGRRRVQ 540  
 DB 481 DTGRVLFEAKLHKGFLTVSSTSTGPPVVVPANGYKFDSSWNQFYS LAPMGTCNGRRRVQ 540

RESULT 5  
 O12305  
 ID O12305 PRELIMINARY; PRT; 542 AA.  
 AC O12305;  
 DT 01-JUL-1997 (TrEMBLrel. 04, Created)  
 DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE SMA capsid protein (Fragment).  
 OS Snow Mountain virus.  
 OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OX NCBI\_TaxID=52276;  
 RN [1]  
 RP SEQUENCE FROM N.A.

```

RC STRAIN=Snow Mountain strain;
RX MEDLINE=97412813; PubMed=9267456;
RA Hardy M.E., Kramer S.F., Treanor J.J., Estes M.K.;
RT "Human calicivirus genotype II capsid sequence diversity revealed by
RT analyses of the prototype Snow Mountain agent.";
RL Arch. Virol. 142:1469-1479 (1997).
RL EMBL; U75682; AAB61685.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR NON_TER 542 542
SQ SEQUENCE 542 AA; 59275 MW; E0C10C4B84DAFDB1 CRC64;

Query Match 83.0%; Score 2361; DB 12; Length 542;
Best Local Similarity 79.3%; Pred. No. 1.5e-178; Indels 8; Gaps 3;
Matches 432; Conservative 57; Mismatches 48;

QY 1 MKMASNDATPSNDGAAGLVPESSNEMALEPVGASLAAPVTGQTNIIIDPWIRTFVQAP 60
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MKMASNDAAAPSTDGAAGLVPESSNEMALEPVGASLAAPVTGQTNIIIDPWIRTFVQAP 60
QY 61 NGEFTVSPRNSPGEILVNLELPELNPYLALARMYNGYAGGMEVQVMLAGNAFTAGKII 120
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 NGEFTVSPRNPAGEVLLNLELPELNPYLALARMYNGYAGGMEVQVMLAGNAFTAGKLV 120
QY 121 FAAPVPFVVENISPSQITMFPHVIIDVRLPEVLLPMDPVRSTLPHFNQKDEPKRLVA 180
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 FAAPVPHFPVENISPSQITMFPHVIIDVRLPEVLLPMDPVRSTLPHFNQKDEPKRIVA 180
QY 181 MLYTPLRNSGGDDVFTVSCRILTRPSDFDFTYLPVPTVESKTKPFTLPIVLTGELS 240
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 MLYTPLRNSGGDDVFTVSCRILTRPSDFDFTYLPVPTVESKTKPFTLPIVLTGELS 240
QY 241 RPLSIDEMWTSNESPESIVQPNQGRVTLDELLGTTQLQACNICSIRGKVTGVPSEQHM 300
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 RFPVSIIDQMTSPNEVISVQCNQGRCTLDGELGTTQLQVSGICAFKGEVTAHLQNDHL 300
QY 301 WNLBITNLNGTQFDPTDDVPAPLGVDPFAGEVGVLSQRNR----GESIPANRAHDAVA 356
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 YNITITNLNGSPDFSEDIAPLGVDPFQGRVGVITQRDKQNAAGSQPANRGHDAVVP 360
QY 357 TYSKYTPKGLVQIGTWTNDVE--NPTKFTPIGLNEVANGHRFEQWTLPIRYSGALTIN 415
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 TYTAQYTPKGLVQIGTWTNDVE--NPTKFTPIGLNEVANGHRFEQWTLPIRYSGALTIN 417
QY 416 MNLAPAVAPLPFGERLLFFRSYVPLKGGFNGNPAIDCSVPQEWVQHFQESAPSLGDVALV 475
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
418 TNLAPSVAPVPFGERLLFFRSYVPLKGGFNGNPAIDCSVPQEWVQHFQESAPSMSEVALV 477
QY 476 RYVNPDTGRVLFEAKLHKGGLTVSSTSTGVPVVPANGYKFKDSWVNOQFYSLAPMGTCNG 535
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
478 RYINPDTGRALFEAKLHKGGLTVSSTSTGVPVVPANGYKFKDSWVNOQFYSLAPMGTCNG 537
QY 536 RRRVQ 540
DB |||||:|||||:
538 RRRIQ 542

RESULT 7
Q96877 PRELIMINARY; PRT; 542 AA.
AC Q96877;
DT 01-FEB-1997 (Tremblrel. 02, Created)
DT 01-FEB-1997 (Tremblrel. 02, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Capsid protein.
OS Snow Mountain virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=52276;
RN [1]
RP SEQUENCE FROM N.A.
RA Lochridge V.P., Hardy M.E.;

"Snow Mountain Virus Genome Sequence and Virus-like Particle Assembly.";
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY134748; AAN08112.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR SEQUENCE 542 AA; 59275 MW; E0C10C4B84DAFDB1 CRC64;

Query Match 83.0%; Score 2361; DB 12; Length 542;
Best Local Similarity 79.3%; Pred. No. 1.5e-178; Indels 8; Gaps 3;
Matches 432; Conservative 57; Mismatches 48;

QY 1 MKMASNDATPSNDGAAGLVPESSNEMALEPVGASLAAPVTGQTNIIIDPWIRTFVQAP 60
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
1 MKMASNDAAAPSTDGAAGLVPESSNEMALEPVGASLAAPVTGQTNIIIDPWIRTFVQAP 60
QY 61 NGEFTVSPRNSPGEILVNLELPELNPYLALARMYNGYAGGMEVQVMLAGNAFTAGKII 120
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
61 NGEFTVSPRNPAGEVLLNLELPELNPYLALARMYNGYAGGMEVQVMLAGNAFTAGKLV 120
QY 121 FAAPVPFVVENISPSQITMFPHVIIDVRLPEVLLPMDPVRSTLPHFNQKDEPKRLVA 180
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
121 FAAPVPHFPVENISPSQITMFPHVIIDVRLPEVLLPMDPVRSTLPHFNQKDEPKRIVA 180
QY 181 MLYTPLRNSGGDDVFTVSCRILTRPSDFDFTYLPVPTVESKTKPFTLPIVLTGELS 240
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
181 MLYTPLRNSGGDDVFTVSCRILTRPSDFDFTYLPVPTVESKTKPFTLPIVLTGELS 240
QY 241 RPLSIDEMWTSNESPESIVQPNQGRVTLDELLGTTQLQACNICSIRGKVTGVPSEQHM 300
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
241 RFPVSIIDQMTSPNEVISVQCNQGRCTLDGELGTTQLQVSGICAFKGEVTAHLQNDHL 300
QY 301 WNLBITNLNGTQFDPTDDVPAPLGVDPFAGEVGVLSQRNR----GESIPANRAHDAVA 356
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
301 YNITITNLNGSPDFSEDIAPLGVDPFQGRVGVITQRDKQNAAGSQPANRGHDAVVP 360
QY 357 TYSKYTPKGLVQIGTWTNDVE--NPTKFTPIGLNEVANGHRFEQWTLPIRYSGALTIN 415
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
361 TYTAQYTPKGLVQIGTWTNDVE--NPTKFTPIGLNEVANGHRFEQWTLPIRYSGALTIN 417
QY 416 MNLAPAVAPLPFGERLLFFRSYVPLKGGFNGNPAIDCSVPQEWVQHFQESAPSLGDVALV 475
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
418 TNLAPSVAPVPFGERLLFFRSYVPLKGGFNGNPAIDCSVPQEWVQHFQESAPSMSEVALV 477
QY 476 RYVNPDTGRVLFEAKLHKGGLTVSSTSTGVPVVPANGYKFKDSWVNOQFYSLAPMGTCNG 535
DB |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
478 RYINPDTGRALFEAKLHKGGLTVSSTSTGVPVVPANGYKFKDSWVNOQFYSLAPMGTCNG 537
QY 536 RRRVQ 540
DB |||||:|||||:
538 RRRIQ 542

RESULT 6
Q80RD6 PRELIMINARY; PRT; 542 AA.
AC Q80RD6;
DT 01-JUN-2003 (Tremblrel. 24, Created)
DT 01-JUN-2003 (Tremblrel. 24, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Viral protein 1.
OS Snow Mountain virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=52276;
RN [1]
RP SEQUENCE FROM N.A.
RA Lochridge V.P., Hardy M.E.;

```



RL Virus Genes 15:5-7(1997).  
DR EMBL: U70059; AAB16915.1; -.  
DR InterPro: IPR004005; Calici\_coat.  
DR InterPro: IPR008975; Viral\_Cap\_coat.  
DR Pfam: PF00915; Calici\_coat; 1.  
SQ SEQUENCE 542 AA; 59201 MW; E7A22C488ABFDB1 CRC64;  
  
Query Match 83.0%; Score 2360; DB 12; Length 542;  
Best Local Similarity 79.1%; Pred. No. 1.8e-178;  
Matches 433; Conservative 56; Mismatches 48; Indels 8; Gaps 3;  
  
QY 1 MKMASNDATPSNDGAGLVPSNNEAMALEPVVGASLAAPVTGQTNIIIDPWIRTFVQAP 60  
DB 1 MKMASNDAAAPSDGAGLVPSNNEAMALEPVVGASLAAPVTGQTNIIIDPWIRTFVQAP 60  
  
QY 61 NGEFTVSPRNSPGEITLVNLELGPENLPYLALHARMYNGYAGMEVQVMLAGNAFTAGKII 120  
DB 61 NGEFTVSPRNPAGFVLLNLELGPENLPYLALHARMYNGYAGMEVQVMLAGNAFTAGKLV 120  
  
QY 121 FFAVPPYFVENVLSQITMFPFHVIDVRLPEVLLPMDPVRSTLTFHFNQKDEPKRLVA 180  
DB 121 FFAVPPHFFVENLSQOITMFPFHVIDVRLPEVLLPMDPVRSTLTFHFNQKDEPKRVI 180  
  
QY 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEFDTYLVPPPTVESKTKPFTLPVLTGELS 240  
DB 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEFDTYLVPPPTVESKTKPFTLPVLTGELS 240  
  
QY 241 RPPLSIDENMTSPNESIVVQPNQGRVTLTGELGTTTQACNICISIRGKVTGQVPSQHM 300  
DB 241 RPVSIDQMTSPNEVISVQCNQGRCTLDGELGTTTQACNICISIRGKVTGQVPSQHM 300  
  
QY 301 WNLEITNLNGTQDPTDDVPAPLGPVDFAGFVGVLSQRNR---GESNPANRAHDVVA 356  
DB 301 YNITITNLNGSPFDSEDIAPLGPVDFQGRVFGVITQRDKQNAAGSQSPANRKHDAVVP 360  
  
QY 357 TYSKXTPKLGLVQIGTWTNDVE-NQPTKFTPIGLNEVANGHRFEQWTLPRYSALTIN 415  
DB 361 TYTAQYTPKLGQVIGTWTDDLVKNQPVKFTFVGLNDT---EHFNQWVVPYAGALNIN 417  
  
QY 416 MNLAPAVAPLPGERLLFRSVPVPLKGGFNPADCSVPQEWQHFYQBSAPSLGDVALV 475  
DB 418 TNLAPSAVPVFPGERLLFRSVPVPLKGGYGNPAIDCLLPQEWQHFYQBAASMSSEVALV 477  
  
QY 476 RYVNPDTGRVLFEAKLHKGFFLTVSSTSGPVVPVANGYFKFDSWVNFYSLAPMGTGNG 535  
DB 478 RYINPDTGRALFEAKLHAGFTVSSNTSAPVVPVANGYFRFDSWVNFYSLAPMGTGNG 537  
  
QY 536 RRRVQ 540  
DB 538 RRRVQ 542

RESULT 8  
Q913B6 PRELIMINARY; PRT; 542 AA.  
ID Q913B6;  
AC Q913B6;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus NUV/Chesterfield/434/1997/US.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=173724;  
RN [1]  
RN SEQUENCE FROM N.A.  
RC STRAIN=NUV/Chesterfield/434/1997/US;  
RX MEDLINE=99034638; PubMed=9815206;  
RA Fankhauser R.L., Noel J.S., Monroe S.S., Ando T., Glass R.I.;  
RT "Molecular epidemiology of 'Norwalk-like viruses' in outbreaks of gastroenteritis in the United States.";  
RL J. Infect. Dis. 178:1571-1578(1998).  
RN [2]

RP SEQUENCE OF 100-192 FROM N.A.  
RC STRAIN=NUV/Chesterfield/434/1997/US;  
RX MEDLINE=99246332; PubMed=10228052;  
RA Noel J.S., Fankhauser R.L., Ando T., Monroe S.S., Glass R.I.;  
RT "Identification of a distinct common strain of 'Norwalk-like viruses' having a global distribution.";  
RT J. Infect. Dis. 179:1334-1344(1999).  
RN [3]  
RN SEQUENCE FROM N.A.  
RC STRAIN=NUV/Chesterfield/434/1997/US;  
RX MEDLINE=20266071; PubMed=10804147;  
RA Ando T., Noel J.S., Fankhauser R.L.;  
RT "Genetic classification of 'Norwalk-like viruses'.";  
RL J. Infect. Dis. 181:S336-S348(2000).  
RN [4]  
RN SEQUENCE FROM N.A.  
RC STRAIN=NUV/Chesterfield/434/1997/US;  
RA Noel J.S., Fankhauser R.L., Monroe S.S., Glass R.I., Ando T.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY054300; AAL15548.1; -.  
DR InterPro: IPR004005; Calici\_coat.  
DR InterPro: IPR008975; Viral\_Cap\_coat.  
DR Pfam: PF00915; Calici\_coat; 1.  
SQ SEQUENCE 542 AA; 59233 MW; 15652994B5631E31 CRC64;  
  
Query Match 82.6%; Score 2351; DB 12; Length 542;  
Best Local Similarity 79.1%; Pred. No. 9.1e-178;  
Matches 431; Conservative 56; Mismatches 50; Indels 8; Gaps 3;  
  
QY 1 MKMASNDATPSNDGAGLVPSNNEAMALEPVVGASLAAPVTGQTNIIIDPWIRTFVQAP 60  
DB 1 MKMASNDAAAPSDGAGLVPSNNEAMALEPVVGASLAAPVTGQTNIIIDPWIRTFVQAP 60  
  
QY 61 NGEFTVSPRNSPGEITLVNLELGPENLPYLALHARMYNGYAGMEVQVMLAGNAFTAGKII 120  
DB 61 NGEFTVSPRNPAGFVLLNLELGPENLPYLALHARMYNGYAGMEVQVMLAGNAFTAGKLV 120  
  
QY 121 FFAVPPYFVENVLSQITMFPFHVIDVRLPEVLLPMDPVRSTLTFHFNQKDEPKRLVA 180  
DB 121 FFAVPPHFFVENLSQOITMFPFHVIDVRLPEVLLPMDPVRSTLTFHFNQKDEPKRVI 180  
  
QY 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEFDTYLVPPPTVESKTKPFTLPVLTGELS 240  
DB 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPEFDTYLVPPPTVESKTKPFTLPVLTGELS 240  
  
QY 241 RPPLSIDENMTSPNESIVVQPNQGRVTLTGELGTTTQACNICISIRGKVTGQVPSQHM 300  
DB 241 RPVSIDQMTSPNEVISVQCNQGRCTLDGELGTTTQACNICISIRGKVTGQVPSQHM 300  
  
QY 301 WNLEITNLNGTQDPTDDVPAPLGPVDFAGFVGVLSQRNR---GESNPANRAHDVVA 356  
DB 301 YNITITNLNGSPFDSEDIAPLGPVDFQGRVFGVITQRDKQNAAGSQSPANRKHDAVVP 360  
  
QY 357 TYSKXTPKLGLVQIGTWTNDVE-NQPTKFTPIGLNEVANGHRFEQWTLPRYSALTIN 415  
DB 361 TYTAQYTPKLGQVIGTWTDDLVKNQPVKFTFVGLNDT---EHFNQWVVPYAGALNIN 417  
  
QY 416 MNLAPAVAPLPGERLLFRSVPVPLKGGFNPADCSVPQEWQHFYQBSAPSLGDVALV 475  
DB 418 TNLAPSAVPVFPGERLLFRSVPVPLKGGYGNPAIDCLLPQEWQHFYQBAASMSSEVALV 477  
  
QY 476 RYVNPDTGRVLFEAKLHKGFFLTVSSTSGPVVPVANGYFKFDSWVNFYSLAPMGTGNG 535  
DB 478 RYINPDTGRALFEAKLHAGFTVSSNTSAPVVPVANGYFRFDSWVNFYSLAPMGTGNG 537  
  
QY 536 RRRVQ 540  
DB 538 RRRVQ 542  
  
RESULT 9  
Q66293 PRELIMINARY; PRT; 542 AA.  
ID Q66293

AC Q66293;  
 DT 01-NOV-1996 (TReMBLrel. 01, Created)  
 DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Capsid protein.  
 OS Human calicivirus NLV/Exfurt/546/00/DB.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae.  
 OC Calicivirus.  
 OX NCBI\_TaxID=11975;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Melksham virus;  
 RX MEDLINE=96136658; PubMed=8533462;  
 RA Green S.M., Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;  
 RT "Capsid diversity in small round-structured viruses : molecular  
 RT characterization of an antigenically distinct human enteric  
 RT calicivirus";  
 RL Virus Res. 37:271-283(1995).  
 DR EMBL; X81879; CAA57462.1; -;  
 DR PIR; S60616; S60616.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 542 AA; 59147 MW; 0A0AA91343E60341 CRC64;  
 Query Match 82.2%; Score 2339; DB 12; Length 542;  
 Best Local Similarity 78.9%; Pred. No. 8.1e-177;  
 Matches 430; Conservative 56; Mismatches 51; Indels 8; Gaps 3;  
 QY 1 MKMASNDATPSNDGAAGLVPESSNEAMALEPVPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60  
 DB 1 MKMASNDATPSNDGAAGLVPESSNEAMALEPVPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60  
 QY 61 NGEFTVSPRNSPGEIILVNLELGPENLPYLAHLARMYNGVAGMEVOMLAGNAFTAGKII 120  
 DB 61 NGEFTVSPRNSPGEIILVNLELGPENLPYLAHLARMYNGVAGMEVOMLAGNAFTAGKII 120  
 QY 121 FAAPVPFPVENISPSQITMFPFHVIIIDVRLPEVLLPMPDVRSTLPHNQDPEKRLVA 180  
 DB 121 FAAPVPFPVENISPSQITMFPFHVIIIDVRLPEVLLPMPDVRSTLPHNQDPEKRLVA 180  
 QY 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPDFDYLVPPVPTVESKTKPFTLPVLTGLSLS 240  
 DB 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPDFDYLVPPVPTVESKTKPFTLPVLTGLSLS 240  
 QY 241 RFPPLSIDEMVTSNPNESIVQFQNGRVTLTGELGTTLQACNICSIRGKVTGQVPSQHM 300  
 DB 241 RFPPLSIDEMVTSNPNESIVQFQNGRVTLTGELGTTLQACNICSIRGKVTGQVPSQHM 300  
 QY 301 WNLLEITNLNGTQFDDTDDVPAPLGPDPFAGEVFGVLSQRNR----GESNPNARAHDAVVA 356  
 DB 301 WNLLEITNLNGTQFDDTDDVPAPLGPDPFAGEVFGVLSQRNR----GESNPNARAHDAVVA 356  
 QY 357 TYSKYTPKGLVQIGTWTNTDVE-NQPTKFTPIGLNEVANGHREFTWTLPRYSGALTIN 415  
 DB 357 TYSKYTPKGLVQIGTWTNTDVE-NQPTKFTPIGLNEVANGHREFTWTLPRYSGALTIN 415  
 QY 416 MNLAPAVAPLPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWQHYQESAPSLGDVALV 475  
 DB 416 MNLAPAVAPLPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWQHYQESAPSLGDVALV 475  
 QY 476 RYVNPDTGRVLFEAKLHKGGLFTVSSSTSTGVPVVPVANGYKFDSSWNQFYSAPMGGTNG 535  
 DB 476 RYVNPDTGRVLFEAKLHKGGLFTVSSSTSTGVPVVPVANGYKFDSSWNQFYSAPMGGTNG 535  
 QY 536 RRRVQ 540  
 DB 536 RRRVQ 542  
 QY 537 RRRVQ 540  
 DB 537 RRRVQ 542  
 RESULT 10  
 Q915C7 PRELIMINARY; PRT; 548 AA.  
 ID Q915C7  
 AC Q915C7;

DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)  
 DE Capsid protein.  
 OS Human calicivirus NLV/Exfurt/546/00/DB.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae.  
 OC Norovirus.  
 OX NCBI\_TaxID=173929;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NLV/Exfurt/546/00/DE;  
 RX Kuenkel U., Hoehne M., Schreier E.;  
 RA "Molecular epidemiology of outbreaks of gastroenteritis associated  
 RT with Norwalk-like viruses in Germany";  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF427118; AAL18874.1; -;  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 548 AA; 59874 MW; 9A1D8BEE58ADE7D CRC64;  
 Query Match 81.8%; Score 2326; DB 12; Length 548;  
 Best Local Similarity 79.2%; Pred. No. 8.9e-176;  
 Matches 435; Conservative 51; Mismatches 53; Indels 10; Gaps 6;  
 QY 1 MKMASNDATPSNDGAAGLVPESSNEAMALEPVPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60  
 DB 1 MKMASNDATPSNDGAAGLVPESSNEAMALEPVPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60  
 QY 61 NGEFTVSPRNSPGEIILVNLELGPENLPYLAHLARMYNGVAGMEVOMLAGNAFTAGKII 120  
 DB 61 NGEFTVSPRNSPGEIILVNLELGPENLPYLAHLARMYNGVAGMEVOMLAGNAFTAGKII 120  
 QY 121 FAAPVPFPVENISPSQITMFPFHVIIIDVRLPEVLLPMPDVRSTLPHNQDPEKRLVA 180  
 DB 121 FAAPVPFPVENISPSQITMFPFHVIIIDVRLPEVLLPMPDVRSTLPHNQDPEKRLVA 180  
 QY 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPDFDYLVPPVPTVESKTKPFTLPVLTGLSLS 240  
 DB 181 MLYTPLRSNGSGDDVFTVSCRILTRPSPDFDYLVPPVPTVESKTKPFTLPVLTGLSLS 240  
 QY 241 RFPPLSIDEMVTSNPNESIVQFQNGRVTLTGELGTTLQACNICSIRGKVTGQVPSQHM- 299  
 DB 241 RFPPLSIDEMVTSNPNESIVQFQNGRVTLTGELGTTLQACNICSIRGKVTGQVPSQHM- 299  
 QY 300 -MNLLEITNLNGTQFDDTDDVPAPLGPDPFAGEVFGVLSQRNR----GESN-PANRAHDA 353  
 DB 301 THWNTVTVNLNGTQFDDTDDVPAPLGPDPFAGEVFGVLSQRNR----GESN-PANRAHDA 360  
 QY 354 VVATYSKYTPKGLVQIGTWTNTDVE-NQPTKFTPIGLNEV-ANGHREFTWTLPRYSGA 411  
 DB 354 VVATYSKYTPKGLVQIGTWTNTDVE-NQPTKFTPIGLNEV-ANGHREFTWTLPRYSGA 411  
 QY 412 LTLNMLNAPAVAPLPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWQHYQESAPSLGD 471  
 DB 420 LTLNMLNAPAVAPLPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWQHYQESAPSLGD 479  
 QY 472 VALVRYNPDTGRVLFEAKLHKGGLFTVSSSTSTGVPVVPVANGYKFDSSWNQFYSAPMG 531  
 DB 480 VALVRYNPDTGRVLFEAKLHKGGLFTVSSSTSTGVPVVPVANGYKFDSSWNQFYSAPMG 539  
 QY 532 TNGRRVQ 540  
 DB 540 TNGRRVQ 548  
 RESULT 11  
 Q911I5 PRELIMINARY; PRT; 535 AA.  
 ID Q911I5  
 AC Q911I5;  
 DT 01-DEC-2001 (TReMBLrel. 19, Created)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)

DE Capsid protein.  
OS Human calicivirus NLV/Schwerin 003/00/DE.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=168852;  
RN [1]\_TaxID=168852;  
RP SEQUENCE FROM N.A.  
RA Kuenkel U., Schreier E.;  
RT "Molecular epidemiology of outbreaks of gastroenteritis associated  
with Norwalk-like viruses in Germany."  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF397905; AK92128.1; -  
DR InterPro: IPR004005; Calici\_coat.  
DR InterPro: IPR008975; Viral\_cap\_coat.  
DR Pfam: PF00915; Calici\_coat\_1.  
SQ SEQUENCE 535 AA; 58520 MW; 2E97BBD3F3D5FCC1 CRC64;

Query Match 79.9%; Score 2273.5; DB 12; Length 535;  
Best Local Similarity 76.5%; Pred. No. 1.2e-171;  
Matches 414; Conservative 65; Mismatches 55; Indels 7; Gaps 3;

QY 1 MKMASNDATPSNDGAGLVPESSNNEAMALEPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60  
DB 1 MKMASNDATPSNDGAGLVPESSNNEAMALEPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60  
QY 61 NGEFTVSPRNSPGEILVNLELPELNPYLALHARMYNGYAGMEVQVMLAGNAFTAGKII 120  
DB 61 NGEFTVSPRNSPGEILVNLELPELNPYLALHARMYNGYAGMEVQVMLAGNAFTAGKII 120  
QY 121 FAAPVPPYFVENLSQITMFPFHVIDVRLPEVLLPMPDVRSTLPHFNQKDEPKRLVA 180  
DB 121 FAAPVPPYFVENLSQITMFPFHVIDVRLPEVLLPMPDVRSTLPHFNQKDEPKRLVA 180  
QY 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDFTYLVPPPTVESKTKPFTLPVLTIGELNS 240  
DB 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDFTYLVPPPTVESKTKPFTLPVLTIGELNS 240  
QY 241 RPPLSIDEMVTSNPSISVVPQNGRVTLDGELLGTTQLQACNICISIRGKVTQVPSQOHM 300  
DB 241 RPPLSIDEMVTSNPSISVVPQNGRVTLDGELLGTTQLQACNICISIRGKVTQVPSQOHM 300  
QY 301 WNLEITNLNGTQFDPTDDVPAPLGVDPFAGEVGVLSQRNRESNPANRAHDVAATYSD 360  
DB 301 WNLEITNLNGTQFDPTDDVPAPLGVDPFAGEVGVLSQRNRESNPANRAHDVAATYSD 360  
QY 361 KYTPKGLGVQIGTWNNDVE--NPTKFTPIGLNEVANGHRFEQWTLPRYSGALTINMILA 419  
DB 361 KYTPKGLGVQIGTWNNDVE--NPTKFTPIGLNEVANGHRFEQWTLPRYSGALTINMILA 419  
QY 420 PAVAPLFFGERLLFFRSYVPLKGGFNPADCSVPQEWVQHFYQBSAPSLGDVALVRYVN 479  
DB 420 PAVAPLFFGERLLFFRSYVPLKGGFNPADCSVPQEWVQHFYQBSAPSLGDVALVRYVN 479  
QY 480 PDTGRVLFEAKLHKGGFTVSTSTGPPVVPANGYFKFSDSWNQYFSLAPMGTCNGRRRV 539  
DB 480 PDTGRVLFEAKLHKGGFTVSTSTGPPVVPANGYFKFSDSWNQYFSLAPMGTCNGRRRV 539  
QY 540 Q 540  
DB 535 Q 535

RESULT 12  
ID Q68104 PRELIMINARY; PRT; 535 AA.  
AC Q68104;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Hawaii calicivirus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.

OX NCBI\_TaxID=33750;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;  
RX MEDLINE=94358485; PubMed=8077710;  
RA Lew J.F., Kapikian A.Z., Valdeusio J., Green K.Y.;  
RT "Molecular characterization of Hawaii virus and other Norwalk-like  
caliciviruses."  
RL J. Infect. Dis. 170:535-542(1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;  
RX Lew J.F.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;  
RX Green K.Y., Sosnovtseva S.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;  
RX Pletneva M.A., Sosnovtseva S., Green K.Y.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: U07611; AAB97768.2; -  
DR InterPro: IPR004005; Calici\_coat.  
DR InterPro: IPR008975; Viral\_cap\_coat.  
DR Pfam: PF00915; Calici\_coat\_1.  
SQ SEQUENCE 535 AA; 58741 MW; DOCA252A210E4599 CRC64;

Query Match 79.8%; Score 2270.5; DB 12; Length 535;  
Best Local Similarity 76.3%; Pred. No. 2.2e-171;  
Matches 413; Conservative 63; Mismatches 58; Indels 7; Gaps 3;

QY 1 MKMASNDATPSNDGAGLVPESSNNEAMALEPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60  
DB 1 MKMASNDATPSNDGAGLVPESSNNEAMALEPVVGASLAAPVTGQTNIIIDPWIRTNFVQAP 60  
QY 61 NGEFTVSPRNSPGEILVNLELPELNPYLALHARMYNGYAGMEVQVMLAGNAFTAGKII 120  
DB 61 NGEFTVSPRNSPGEILVNLELPELNPYLALHARMYNGYAGMEVQVMLAGNAFTAGKII 120  
QY 121 FAAPVPPYFVENLSQITMFPFHVIDVRLPEVLLPMPDVRSTLPHFNQKDEPKRLVA 180  
DB 121 FAAPVPPYFVENLSQITMFPFHVIDVRLPEVLLPMPDVRSTLPHFNQKDEPKRLVA 180  
QY 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDFTYLVPPPTVESKTKPFTLPVLTIGELNS 240  
DB 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDFTYLVPPPTVESKTKPFTLPVLTIGELNS 240  
QY 241 RPPLSIDEMVTSNPSISVVPQNGRVTLDGELLGTTQLQACNICISIRGKVTQVPSQOHM 300  
DB 241 RPPLSIDEMVTSNPSISVVPQNGRVTLDGELLGTTQLQACNICISIRGKVTQVPSQOHM 300  
QY 301 WNLEITNLNGTQFDPTDDVPAPLGVDPFAGEVGVLSQRNRESNPANRAHDVAATYSD 360  
DB 301 WNLEITNLNGTQFDPTDDVPAPLGVDPFAGEVGVLSQRNRESNPANRAHDVAATYSD 360  
QY 361 KYTPKGLGVQIGTWNNDVE--NPTKFTPIGLNEVANGHRFEQWTLPRYSGALTINMILA 419  
DB 361 KYTPKGLGVQIGTWNNDVE--NPTKFTPIGLNEVANGHRFEQWTLPRYSGALTINMILA 419  
QY 420 PAVAPLFFGERLLFFRSYVPLKGGFNPADCSVPQEWVQHFYQBSAPSLGDVALVRYVN 479  
DB 420 PAVAPLFFGERLLFFRSYVPLKGGFNPADCSVPQEWVQHFYQBSAPSLGDVALVRYVN 479  
QY 480 PDTGRVLFEAKLHKGGFTVSTSTGPPVVPANGYFKFSDSWNQYFSLAPMGTCNGRRRV 539  
DB 480 PDTGRVLFEAKLHKGGFTVSTSTGPPVVPANGYFKFSDSWNQYFSLAPMGTCNGRRRV 539  
QY 540 Q 540  
DB 535 Q 535

Db 535 Q 535

RESULT 13

Q9QMK6 PRELIMINARY; PRT; 535 AA.

AC Q9QMK6; (1)

DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)

DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)

DE Capsid protein.

OS Chitt virus.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.

OC NCBI\_TaxID=104901;

OX NCBI\_TaxID=104901;

RN (1)

RP SEQUENCE FROM N.A.

RC STRAIN=Aichi76-96;

RX MEDLINE=20471382; PubMed=11021399;

RA Kobayashi S., Sakae K., Suzuki Y., Ishiko H., Kamata K., Suzuki K., Natori K., Miyamura T., Takeda N.;

RT "Expression of recombinant capsid proteins of chitt virus, a genogroup II Norwalk virus, and development of an ELISA to detect the viral antigen.;"

RL Microbiol. Immunol. 44:687-693 (2000).

DR EMBL; AB032758; BAA84716.1; -.

DR InterPro; IPR004005; Calici.coat.

DR InterPro; IPR008975; Viral.Cap.coat.

DR Pfam; PF00915; Calici.coat; 1.

SQ SEQUENCE 535 AA; 58504 MW; 5CCEBC91BFC7B8B CRC64;

Query Match 79.8%; Score 2269.5; DB 12; Length 535;

Best Local Similarity 76.3%; Pred. No. 2.6e-171;

Matches 413; Conservative 65; Mismatches 56; Indels 7; Gaps 3;

Qy 1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVGASIAAPVTGQNIIDPWIRTNFVQAP 60

Db 1 MKMASNDAAAPSNDGAAGLVPEANNETWALEPVGASIAAPLTQNNIIDPWIRTNFVQAP 60

Qy 61 NGEFTVSPRNSPGEILVNLGPELNPYLAHLARMYNGYAGGVEQVLLAGNAFTAGKII 120

Db 61 NGEFTVSPRNSPGEVGLNLELGPENLPYLAHLARMYNGYAGGVEQVLLAGNAFTAGKLV 120

Qy 121 FAAPVPPFVLENLSQITMPPHVIIDVRLTPVLLPMPDVRSTLPHFNQKDEPKMLVA 180

Db 121 FAAPVPPFLENISPGQITMPPHVIIDVRLTPVLLPMPDVRNFFHYNQNEPRMLVA 180

Qy 181 MLYTPLRNSGDDVFTVSCRILTRPSPDFVLPPTVESKTKPFTLPLVLTGELSNS 240

Db 181 MLYTPLRNSGDDVFTVSCRILTRPSPDFVLPPTVESKTKPFTLPLIITGELTNS 240

Qy 241 RPLSLIDEMVTSNBSIVVQPNQGRVTLDGELGTQLQACNICISIRGKVTGQVPSEQHM 300

Db 241 MLYTPLRNSGDDVFTVSCRILTRPSPDFVLPPTVESKTKPFTLPLIITGELTNS 240

Qy 241 RPLSLIDEMVTSNBSIVVQPNQGRVTLDGELGTQLQACNICISIRGKVTGQVPSEQHM 300

Db 241 RFPVPIDELYTSPNESLVVQPNQGRCALDGLQGTQLLPTAICSFGRINQKVSNGH 300

Qy 301 WNLITNLNGTQFDPTDDVPAPLGVPDFAGEVFGVLSQRNGESNPANRAHDAVATYSD 360

Db 301 WNLITNLNGTQFDPTDDVPAPLGVPDFAGEVFGVLSQRNGESNPANRAHDAVATYSD 360

Qy 301 WNLITNLNGTQFDPTDDVPAPLGVPDFAGEVFGVLSQRNGESNPANRAHDAVATYSD 360

Db 301 WNLITNLNGTQFDPTDDVPAPLGVPDFAGEVFGVLSQRNGESNPANRAHDAVATYSD 360

Qy 361 KYTEPKLGLVQIGTWNNDVE-NQPTKFTPIGLINEVANGHRFEQWTLPRYSGALTLNNMLA 419

Db 361 KYTEPKLGLVQIGTWNNDVE-NQPTKFTPIGLINEVANGHRFEQWTLPRYSGALTLNNMLA 419

Qy 361 KYTEPKLGLVQIGTWNNDVE-NQPTKFTPIGLINEVANGHRFEQWTLPRYSGALTLNNMLA 419

Db 361 KYTEPKLGLVQIGTWNNDVE-NQPTKFTPIGLINEVANGHRFEQWTLPRYSGALTLNNMLA 419

Qy 420 PAVAPLPFGERLLFFRSYVPLKGGFGNPAIDCSVPOEWQHFYQESAPSLGDVALVRYN 479

Db 420 PAVAPLPFGERLLFFRSYVPLKGGFGNPAIDCSVPOEWQHFYQESAPSLGDVALVRYN 479

Qy 420 PAVAPLPFGERLLFFRSYVPLKGGFGNPAIDCSVPOEWQHFYQESAPSLGDVALVRYN 479

Db 420 PAVAPLPFGERLLFFRSYVPLKGGFGNPAIDCSVPOEWQHFYQESAPSLGDVALVRYN 479

Qy 479 PVTGRLVFEAKLHKGGLTVSSTSTGTVVVPVANGYFKFDSWVNFYSLAPMGTGNGRRV 539

Db 479 PVTGRLVFEAKLHKGGLTVSSTSTGTVVVPVANGYFKFDSWVNFYSLAPMGTGNGRRV 539

Qy 540 Q 540

Db 535 Q 535

```
RESULT 15
Q917X7 PRELIMINARY; PRT; 535 AA.
AC Q917X7;
DC 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Honolulu/314/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171852;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Honolulu/314/1994/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Honolulu/314/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
characterized small round-structured viruses involved in outbreaks of
nonbacterial acute gastroenteritis in the United States, 1990 to
1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Honolulu/314/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Honolulu/314/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414420; AAL13013.1; --
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58519 MW; C3B7A70CF3617F71 CRC64;

Query Match 79.7%; Score 2268.5; DB 12; Length 535;
Best Local Similarity 76.3%; Pred No. 3.1e-171;
Matches 413; Conservative 65; Mismatches 56; Indels 7; Gaps 3;

QY 1 MKMASNDATPSNDGAGLVPESSNEAMALEPVVVGASLAAPVTGQNTIIDPWIRTNFVQAP 60
DB 1 MKMASNDAAAPSNDGAGLVPEANNEETMALEPVVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
QY 61 NGEFTVSPRNSGCEILVNLLELGPENLPYLALARMYNGYAGGVEQVLMAGNFTAGKII 120
DB 61 NGEFTVSPRNSGCEILVNLLELGPENLPYLALARMYNGYAGGVEQVLLAGNFTAGKLV 120
QY 121 FAAPVPPYPPVENLSQITMPPHVIDVRLTPVLLPMPDVRSLTFHFNQDKPKMRLVA 180
DB 121 FAAPVPPHPPLENISQGIAMFPFHVIDVRLTPVLLPMPDVRNFFHYNQQNEPRRLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDTYLVPPVTESKTKPFTLPVLTLGELNS 240
DB 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDTYLVPPVTESKTKPFTLPVLTLGELNS 240
QY 241 RPPLSIDEMVTSNPSISVVQPONGRVTLTGELLGTQLOACNICIRGKVTGQVPSQHM 300
DB 241 RPPLSIDEMVTSNPSISVVQPONGRVTLTGELLGTQLOACNICIRGKVTGQVPSQHM 300

Db 241 RPPVPIDELYTSPNESLVVQPONGRCVLDGELQGTQLLPTAICSFGRINOKVSGENVH 300
QY 301 WNLEITNLNGTQDPDTPAPPLGVPDFAGVYFVGLSQNRNGESNPANRAHDAVVATYSD 360
DB 301 WNMQVTNINGTPDPTDPAPPLGTPDFSGKLVGLSOR---DHDNACRSHDAVIATNSA 357
QY 361 KYTPKGLGLVOIGTWTNVDVE-NQPTKFTPIGLNEVANGHRFPQWTLPRYSGALTLMNLA 419
DB 358 KFTPKLGAIGTWEEDDVHINQPTKFTPVGLFE---NEGPNQWTLPNYSGALTLMNGLA 414
QY 420 PAVAPLFPGERLLPFRSYVPLKGFNGPAIDCSVPQEWVQHFYQESAPSLGDVALRVYN 479
DB 415 PPVAPTFFQEQQLPFRSHIPLKGGVADPVIDCLLPQEWIQHLYQESAFSQSDVALIRFTN 474
QY 480 PDTGRVLFEAKLHKGGFLTVSSTSTGPPVVVPANGYFKFDSWYNQFYSLAPMGTGNGRRRV 539
DB 475 PDTGRVLFEAKLHRSYITVANTGSRPIVVPANGYFRFDSWYNQFYSLAPMGTGNGRRRV 534
QY 540 Q 540
DB 535 Q 535

Search completed: June 1, 2004, 13:53:19
Job time : 33.4396 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.323 Seconds  
(without alignments)  
4985.230 Million cell updates/sec

Title: US-09-926-799-8  
Perfect score: 2843  
Sequence: 1 MKMASNDAPSDNGAAGLVP.....VNQFVSLAPMTGNGRRRVQ 535  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2298.5	80.8	542	2 S60616	capsid protein - h
2	1923	67.6	539	2 S40111	capsid protein - h
3	1196.5	42.1	530	2 B37471	capsid protein - N
4	1174	41.3	546	2 B37491	major capsid prote
5	328	11.5	576	2 A53982	capsid protein - E
6	327.5	11.5	2344	2 S55399	genome polyprotein
7	322.5	11.3	2344	1 RRWRH	genome polyprotein
8	314.5	11.1	2344	2 S64740	genome polyprotein
9	289	10.2	702	2 A48562	coat protein - San
10	284	10.0	668	1 VCMWFC	coat protein - fel
11	277.5	9.8	668	2 JQ2354	capsid protein - f
12	277.5	9.8	671	1 VCMWF9	coat protein - fel
13	273.5	9.6	703	1 C48562	coat protein - San
14	272.5	9.6	668	1 VCMWFF	coat protein - fel
15	266	9.4	668	2 JQ2356	capsid protein - f
16	130	4.6	2206	1 GNNY27	genome polyprotein
17	128	4.5	2194	1 GNNY7	genome polyprotein
18	127	4.5	733	2 JQ1892	capsid protein - f
19	126	4.4	733	2 JQ1891	capsid protein - f
20	125.5	4.4	2333	1 GNNY2F	genome polyprotein
21	122	4.3	378	2 E96787	protein T402.5 [i
22	121.5	4.3	2332	1 GNNYF	genome polyprotein
23	120.5	4.2	2205	1 GNNY2W	genome polyprotein
24	120	4.2	1011	1 GNNYC1	genome polyprotein
25	119.5	4.2	2336	2 S37077	genome polyprotein
26	115.5	4.1	757	2 JS0198	genome polyprotein
27	115	4.0	2207	2 S09553	genome polyprotein
28	115	4.0	2332	1 GNNY4F	genome polyprotein
29	114	4.0	2206	2 S03822	genome polyprotein

30	114	4.0	2207	1 GNNY5P	genome polyprotein
31	112	3.9	2206	1 GNNY4P	genome polyprotein
32	110.5	3.9	2468	2 AB3412	hypothetical prote
33	110	3.9	716	2 S45262	NF-AT component -
34	109	3.8	833	2 S45041	genome polyprotein
35	109	3.8	913	2 JG0168	gob-5 protein - mo
36	109	3.8	1072	2 S76888	hypothetical prote
37	109	3.8	2201	1 GNNYA9	genome polyprotein
38	108.5	3.8	3262	2 AH2137	hypothetical prote
39	107.5	3.8	929	2 A44048	genome polyprotein
40	107	3.8	703	2 H86349	protein F8K7.9 [im
41	106.5	3.7	3085	2 T00327	polyprotein - infe
42	102.5	3.6	800	2 H84740	hypothetical prote
43	102.5	3.6	1698	2 T13800	coracle gene prote
44	102	3.6	2629	2 I46569	apolipoprotein B -
45	101.5	3.6	852	1 GNNYHA	genome polyprotein

ALIGNMENTS

RESULT 1

S60616  
capsid protein - human calicivirus (strain Melksham)

C:Species: human calicivirus

A:Variety: strain Melksham

C:Date: 23-May-1997 #sequence\_revision 23-May-1997 #text\_change 28-Jul-2000

C:Accession: S60616

R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.

Virus Res. 37, 271-283, 1995

A:Title: Capsid diversity in small round-structured viruses: molecular characterization

A:Reference number: S60615; MUID:96136658; PMID:9533462

A:Accession: S60616

A:Molecule-type: genomic RNA

A:Residues: 1-542 <GRE>

A:Cross-references: ENBL:X81879; NID:G976077; PIDN:CAA57462.1; PID:G976079

A:Experimental source: strain Melksham

A:Note: it is uncertain whether Met-1 or Met-3 is the Initiator

C:Superfamily: human calicivirus capsid protein

C:Keywords: capsid protein; coat protein

Query Match 80.8%; Score 2298.5; DB 2; Length 542;

Best Local Similarity 76.6%; Pred. No. 7.3e-162;

Matches 415; Conservative 65; Mismatches 55;

Indels 7; Gaps 1;

QY 1 MKMASNDAPSDNGAAGLVPPEANNETMALEPVAGASIAAPLTGONNIIDPWIRLNFVQAP 60

Db 1 MKMASNDAPSDTGAAGLVPSNNEVMALEPVAGAAALAPVTGQTNIIDPWIRLNFVQAP 60

QY 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMYNGVAGGVEVQVLLAGNAFTAGKLV 120

Db 61 NGEFTVSPRNAPGEVLLNLELGPENLPYLAHLARMYNGVAGGMEVQVMLAGNAFTAGKLV 120

QY 121 FAAPVPHPLENISQOITMPEPHVIDVRLPLEVLLPDVNRNFFHYNQNEPRMLVA 180

Db 121 FAAPVPHFPVENLSQKITMPEPHVIDVRLPLEVLLPDVNRNFFHYNQKDDPKRIVVA 180

QY 181 MLYTLRNSGSGDDVFTVSCVRLTRPSDFDFNLYLPPTLESKTPFTLPITIGELTNS 240

Db 181 MLYTLRNSGSGDDVFTVSCVRLTRPSDFDFNLYLPPTLESKTPFTLPITIGELTNS 240

QY 241 RFPVPEIDELTSPNBSLVVQPNRGALDGLQGTQLLPTAICSFGRINOKVSGENHV 300

Db 241 RFPVPEIDQMTSPNEVISVQCNQGRCTLDGELQGTQLQVSGICAFKGEVTAHLHDNDHL 300

QY 301 WNMQVNLNGPPFDPTGDPAPLPTDFSGKLFVLSORDHDNAC-----RSHDAVTA 353

Db 301 NNVTITNLNGSPFDPSDIPAPLGVDFQGRVFGVISQRDKQNAAGHSEPNRGRHDVVP 360

QY 354 TNSAKFTPLGAIQIGTWEEDDVHNQPTKFTPVGLFENEGFNQWTLPNYSALTINMGL 413

Db 361 TYTAQYTKLGGIQIGTWQTDLDLTVNQPKFTPVGLNDEHFNQWVPRVYAGALNLTNL 420





A:Reference number: A37491; MUID:93142023; PMID:8380940

A:Accession: B37491

A>Status: not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-546 <LAM>

A:Cross-references: GB:L07418; NID:q1236787; PIDN:AAA92984.1; PID:g2951114

A>Note: sequence extracted from NCBI backbone (NCBI:123456)

A:Title: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype

C:Superfamily: human calicivirus capsid protein

C:Keywords: glycoprotein

F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 41.3%; Score 1174; DB 2; Length 546;

Best Local Similarity 46.2%; Pred. No. 7.6e-79;

Matches 261; Conservative 82; Mismatches 172; Indels 50; Gaps 16;

QY 1 MMASNDAAPSNDGAAG--LVPEANN-ETMALEPVAGASIAAPLTGQNNIIDPWLRLNF 56

DB 1 MMASKDAPOSADGASGAGQLVPEVNTADPLMEPVAGTTAVATAGQNNIDPVLVNF 60

QY 57 VOAPNGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMYNGYAGGVEQVLLAGNAFTA 116

DB 61 VQSPQGEFTISPNTPGDIPLDLGLHLPNPLSHLSQMYNGWGNMVRILLAGNAFSA 120

QY 117 GKLVFAAVPPHPLENIISPGQITMFPVHVIDVRLPVLPLPDVRRNPFPHYNOQNEPRM 176

DB 121 GKIIIVCCVPPGPTSSSLTAQATLPHVIAVRTLEPIEMPLEDVRNLYHTN-DNQPTM 179

QY 177 RLAVMLYPLRS-NGSGD-DVFTVSCRVLTRSPDPDFNLYVPPTLESKTKPFTLPILT 234

DB 180 RLVCMLYPLRTGGSGNSDSFVAVAGRLTAPSSDFSLFVPPTEOKTRAFVFNIP 239

QY 235 GELTNSRFPVPIDELYTSPNESLVVQPNQRCALDGLQTTQLLPTAICSPGRINQKV 294

DB 240 QTLNSRFPVPIDELYTSPNESLVVQPNQRCALDGLQTTQLLPTAICSPGRINQKV 299

QY 295 SCENHVMNQVNTINGTPDTPGDVAPLPTDPFGKLVLSQRDHNAKSHDAV--- 351

DB 300 R-----TLNTEVDGKPF-MAFDSAPVGFDF-GKCDWMHMRISKTNNCTSGDPMRSV 351

QY 352 -IATNSAKFTPKLGAIO-----IGTWEEDDVHINQPTKFTVGLFENEGFNQW 398

DB 352 SVQTNVQGFVPHLSIQDFEVENHTPDYIGFTIE-----WISQPS--TPPG-----TDINLW 401

QY 399 TLPNYSGALTLMNGLAPPVAPTFPGEQILFRSHIPLKGGVADP-VIDCLLPQEWIOHLY 457

DB 402 ELPDYGSSLSQANLAPPVFPFGCEALVYFVSAPFGPNRSAPNDVCLLPQEWIHFV 461

QY 458 QESAPSDVALIRFTNPDTPGRVLPEAKLHRSYITVA--NTGSRPIVVPANGYFRFTW 515

DB 462 SEQAPTMGDAALLHVDPDTNRNLGEFKLYPGGYLTCVPNGVGAGPQQLPLNGVFLFSW 521

QY 516 VNOFYSLAPMGTCN-----GRRV 534

DB 522 VSRFYQLKPVGTPASTARGELGVRI 546

#### RESULT 5

A:33982

capsid protein - European brown hare syndrome virus

C:Species: European brown hare syndrome virus

C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 28-Jul-2000

C:Accession: A53982

R:Wirblich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Ekenas, U.; Haas, B.; Thiel, H.

J. Virol. 68, 5164-5173, 1994

A:Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease

A:Reference number: A53982; MUID:94309183; PMID:7518531

A:Accession: A53982

A:Molecule type: mRNA

A:Residues: 1-576 <WIR>

A:Cross-references: GB:U09199

C:Superfamily: human calicivirus capsid protein

Query Match

Best Local Similarity 11.5%; Score 328; DB 2; Length 576;

Matches 151; Conservative 77; Mismatches 160; Indels 186; Gaps 35;

QY 3 MASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWLRLNFVQANG 62

DB 31 VASTDVVTADNVAAS-----VATAGIGGP-PQASPOESW-RVNFY--ND 72

QY 63 EFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMYNGYAGGVEQVLLAGNAFTAGLVPA 122

DB 73 VFTWSVTDAPOSILYVQHSQNNPFTQVLSQMYAGWAGMQFRIVAGSGIFGRLVCA 132

QY 123 AVPHFPLENTSPG-QITMFPVHVIDVRLPVLPLPDVRRNPFPHYNOQNEPRM--RLV 179

DB 133 IIPPGI---QIQGLEVRQFPHVVIDARSLEPVTITMPDLRPEMYH--PTGDPGLVPTLV 187

QY 180 AMLYTPLRNSGDDVFTVSCRVLTRSPDPDFNLYVPPTLESKT-----KP---ETLPL 232

DB 188 VSVNNL-INPFGGTTSAIQVTETRPSEDFEFLIRAPS--SKTVDSVNSWLLTTPVL 244

QY 233 TIGELTNSRFPVPIDELYTSPNESLVVQPNQRCALDGLQTTQLLPTAICSPGRINQ 292

DB 245 T-GAGSDNRWAPI-----VGLQVPVG----- 265

QY 293 KVSNGHVMNQVNTINGTPF-----DPTGDVPAPLPTDPFGKLVLSQRDH 341

DB 266 GFSTSNRHHNM-----NGSTYGSPPRDDIDHPSGNVSYPTGSATNIETW-----Y 313

QY 342 DNACSHDAVATNSAKFTPKLGAIOI-GTWEEDDVHINQPTKFTVGLFENEGFNQ-WT 399

DB 314 ANAGTATTNPISNTAPDGFDPMGAIPFSGT-----TIPTGAW--VGFGQVWN 358

QY 400 LPN---YSGAL-TLNMGLA-----PPVATFPGEQI-----LPFR 430

DB 359 ASNGTPYVGTQVAVELGANGAPSSIRPVTTT-TGAQLVAKSIYCVATAQNSSAGIIFL 417

QY 431 SHILPKGADVPI--DCLLPQEWIOHLYQSSAPSDVALIRFTNPD-----GR--- 479

DB 418 S---KGMVSTPGVAATYTFQ-----PS---AIV--TTPGTTPVAAPIGKNT 456

QY 480 VLPEAKLHRSYI-----TVANTGSRPIV 504

DB 457 IMFSAVVRTGDVNAVGPVNGVQVGVGSQPLSV 490

#### RESULT 6

S55399

genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)

A:Species: rabbit hemorrhagic disease virus

A:Variety: isolate BS89

C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 18-Jun-1999

C:Accession: S55399

R:Roessli, C.

submitted to the EMBL Data Library, May 1995

A:Reference number: S55399

A:Accession: S55399

A:Molecule type: genomic RNA

A:Residues: 1-2344 <ROS>

A:Cross-references: EMBL:X87607; NID:9854640; PIDN:CAA60910.1; PID:9854641

A:Experimental source: isolate BS89

C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase

C:Keywords: polyprotein

Query Match

Best Local Similarity 11.5%; Score 327.5; DB 2; Length 2344;

Matches 151; Conservative 74; Mismatches 226; Indels 125; Gaps 29;

QY 9 APSNDGAAGLVPEAN---NETMALEP-----VAGASIAAPLTQNNII 48

DB 1773 APOGE-AAGTATTASVPGTTTDDGLDGPVATTSSVVAENSSASITAGIGP-PQVQDQ 1830

QY 49 DPWIRLNFVQAPNGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMYNGYAGGVEQVL 108

1831 ETV-RTNFY--NDVFTWSVADAGSILYTVQHSFQNNPFTAVLSQYAGWAGMQPRFI 1887  
 109 LAGNAFTAGKLVFAAAPPFFLENIISPG-QITMPPHVIIDVRLTLEPVLPLPVRNFFH 167  
 1888 VAGSGVFGRLVAAPVPGI--EIGPGLVRQPHVVIDARSLEPVTITMPDLRPNMYH 1944  
 168 YNQONEPRM--RLVAMLTPLRSNGSGDDVFTVSCRVLTRPSDPDFENYLVPPLESKT- 224  
 1945 --PTGDPGLVPTLVLSVNNL--INPFGSGTSAIQVTVTRESSEDEFWIRAPS--SKTV 1959  
 225 ----KP---FTLPILITIGELTNSRFPVIDELTSPNESLV-VQPONGCA-----LDG 270  
 2000 DSISPAGLLTTPVLT-GVGNDNRW-----NGQIVGLQPVPGFSTCNRHWNLN 2047  
 271 ELQGTQLLPTAICSPFRGINOKVGENHVMQVNTNINPTDPTGDPVAPLGTDPFSG 330  
 2048 STYGWSSPRFADIDHRRGSASYPGSNATNVLOFWYANAGSVDNFIQV-APDGFPMDSF 2106  
 331 KLPGLVLSQRDHDAVATNSAKFTPKLGAIQGT-WEEDDVHINOPT-KFTPVG 388  
 2107 VPF-----NGPGIPAAAGWVGFGAIWNSNSGAPNVTTVQAYELG 2144  
 389 LPENEGFNQWLTLPNYSQALTLNMLAPPVAPTPFGEQILFFRSHIPLKGGVADPVIDCLL 448  
 2145 FATGAPGNLQPTTNTSGAQTVAKSIYAVVTGTAGNPAGLF-----VNASGIIISTPN 2200  
 449 ----PQEWIQLHYSAPSDVALIRFTNPDTRGVLFEAKLHRSYI-----TVA 495  
 2201 YTPQPDRIV-----TTPGTAAAPVGNKNTP-----IMFASVVRRTGVDNATAGSANGTQY 2250  
 496 NTGSRPTW----PANGYF-----RFDTWVNFQFYS 521  
 2251 GTGSQLPVTIGLSLNNYSALMPGQFFVWQLTFAS 2286

RESULT 7  
 RRRWRH  
 genome polyprotein - rabbit hemorrhagic disease virus  
 N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
 C:Species: rabbit hemorrhagic disease virus  
 C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998  
 R:Meiyeers, G.; Wirblich, C.; Thiel, H.J.  
 A:Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing  
 A:Reference number: A41039; MUID:91361557; PMID:1840711  
 A:Accession: A41039  
 A:Molecule type: genomic RNA  
 A:Residues: 1-2344 <MEY>  
 A:Cross-references: GB:M67473  
 C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
 C:Keywords: nucleotidyltransferase

Query Match 11.3%; Score 322.5; DB 1; Length 2344;  
 Best Local Similarity 26.5%; Pred. No. 4.5e-15;  
 Matches 154; Conservative 72; Mismatches 221; Indels 135; Gaps 31;

8 AAPSDGAAGLVPEAN--NETMALEP-----VAGASIAAPLTGQNNI 47  
 1772 AAPQGE-AAGTATTASVPGTTTDMGMDPGVVATTVSITAENSASATATAGIGP-PQQVDQ 1829  
 48 IDPWIRLNFVQAPNGEFTVSPNSGCEVLNLELGPENLPYLAHLRMYNGYAGGVEVQV 107  
 1830 QETW-RTNFY--NDVFTWSVADAGSILYTVQHSFQNNPFTAVLSQYAGWAGMQPRF 1886  
 108 LLAGNAFTAGKLVFAAAPPFFLENIISPG-QITMPPHVIIDVRLTLEPVLPLPVRNFFH 166  
 1887 IVAGSGVFGRLVRAVIPPQI--EIGPGLVRQPHVVIDARSLEPVTITMPDLRPNMY 1943  
 167 HYNQONEPRM--RLVAMLTPLRSNGSGDDVFTVSCRVLTRPSDPDFENYLVPPLESKT 224  
 1944 H--PTGDPGLVPTLVLSVNNL--INPFGSGTSAIQVTVTRESSEDEFWIRAPS--SKT 1998

225 ----KP---FTLPILITIGELTNSRFPVIDELTSPNESLV-VQPONG-----RCALDGELO 273  
 1999 VDSISPAGLLTTPVLT-GVGNDNRW-----NGQIVGLQPVPGFSTCNRHWNLN 2046  
 274 GTTQLLPTA-----ICSPFRGINOKVGENHVMQVNTNINPTDPTGDPVAPLGTDPFS 329  
 2047 GSTYQWSSPRFADIDHRRGSASYPGSNATNVLOFWYANAGSVDNFIQV-APDGFPMDS 2105  
 330 GKLFGVLRSQRDHDAVATNSAKFTPKLGAIQGT-WEEDDVHINOPT-KFTPV 387  
 2106 FVPF-----NGPGIPAAAGWVGFGAIWNSNSGAPNVTTVQAYEL 2143  
 388 GLFENEGFNQWLTLPNYSQALTLNMLAPPVAPTPFGEQILFFRSHIPLKGGVADPVIDCL 447  
 2144 GFATGAPGNLQPTTNTSGAQTVAKSIYAVVTGTAGNPAGLF-----VNASGIIISTPN 2195  
 448 LPOEWIQLHYSAPSDVALIRFTNPD-----GR-----VLPFAKLHRSYI----- 492  
 2196 ----AQRITYTPQPDRI-----VTPGTAAAPVGNKNTPIMFASVVRRTGVDNATAGS 2244  
 493 ----TVANTGSRPTW----PANGYF-----RFDTWVNFQFYS 521  
 2245 ANGTOYGTGSQLPVTIGLSLNNYSALMPGQFFVWQLTFAS 2286

RESULT 8  
 S64740  
 genome polyprotein - rabbit hemorrhagic disease virus (isolate AS7/89)  
 N:Contains: VP60 protein  
 C:Species: rabbit hemorrhagic disease virus  
 A:Variety: isolate AS7/89  
 C:Date: 12-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 20-Jun-2000  
 C:Accession: S64740; S46944; S49018; S65012  
 R:Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.  
 A:Submitted to the EMBL Data Library, May 1995  
 A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by di  
 A:Reference number: S64740  
 A:Accession: S64740  
 A:Molecule type: genomic RNA  
 A:Residues: 1-2344 <CAS>  
 A:Cross-references: EMBL:249271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552  
 A:Experimental source: isolate AS7/89  
 R:Boga, J.; Casais, R.; Martin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.  
 A:Submitted to the EMBL Data Library, July 1993  
 A:Description: Molecular cloning, sequence and expression of the capsid protein gene fr  
 A:Reference number: S46944  
 A:Accession: S46944  
 A:Molecule type: genomic RNA  
 A:Residues: 1650-2344 <BOG>  
 A:Cross-references: EMBL:224757; NID:g515622; PIDN:CAA80881.1; PID:g515623  
 R:Parra, F.; Boga, J.A.; Martin, M.S.; Casais, R.  
 A:Virus Res. 27, 219-228, 1993  
 A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp  
 A:Reference number: S49018; MUID:93255896; PMID:8488721  
 A:Accession: S49018  
 A:Molecule type: genomic RNA  
 A:Residues: 1650-1796 <PAR>  
 A:Cross-references: EMBL:224757  
 A:Experimental source: isolate AS7/89  
 A:Accession: S65012  
 A:Molecule type: protein  
 A:Residues: 1767-1779; 1875-1877; 'X', 1879-1881; 1936-1938; 'X', 1940-1941 <PAW>  
 C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
 C:Keywords: blocked amino end; polyprotein

Query Match 11.1%; Score 314.5; DB 2; Length 2344;  
 Best Local Similarity 25.9%; Pred. No. 1.8e-14;  
 Matches 149; Conservative 75; Mismatches 227; Indels 125; Gaps 30;

9 APSNDGAAGLVPEAN--NETMALEP-----VAGASIAAPLTGQNNI 48



394 TISEKNGSKLIGVATDVIIIGIPDGDWPTTIAADKLIPAGDYSITTEGNDIKTAQAYDT 453  
325 -----TDFSGKLFVLSQRDHNDACHSHDAVIAT-----NSAKFTPKLGAIOIGTWEE 373  
454 AAVVNTTFRGMYICGSLQRAWGDKKISNTAFITATRDGNEIKPSNTIDMTKLAVYQ- 512  
374 DDVHINOPTK-----FTPVGLFENEGFNQWTLNYSGLATLNNMGLAPPVAPFPFG 423  
513 -DTHVEQEVQTSDDTLALLGYTGIG-EAAGSNRRDVRVIS-----VLPAGARGG 561  
424 EQILFRSHIPL 435  
562 NHPFVKNISKL 573

RESULT 11  
JQ2354  
capsid protein - feline calicivirus (strain NADC)  
C:Species: feline calicivirus  
C>Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
C:Accession: JQ2354  
R;Seal, B.S.; Riopach, J.F.; Mengeling, W.L.  
J. Gen. Virol. 74, 2519-2524, 1993  
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable  
A:Reference number: JQ2354; MUID:94065683; PMID:7504075  
A:Accession: JQ2354  
A:Molecule type: mRNA  
A:Residues: 1-668 <SEA>  
A:Cross-references: GB:L09718; NID:g305104; PIDN:AAA16485.1; PID:g305105  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein

Query Match 9.8%; Score 277.5; DB 2; Length 668;  
Best Local Similarity 32.1%; Pred. No. 1.6e-12;  
Matches 79; Conservative 35; Mismatches 93; Indels 39; Gaps 9;

30 EPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPGVGLLNLELGPENLPY 88  
145 EPSQMTAADMATGKSVDSWEAFPSFTSVNW-----STSETQGKILFKQSLGPLNLY 200  
89 LAHLRMYNGVAGGVEVQVLLAGNAFTAGKLVFAAAPPHPPLENISPGQIT---MPPHVI 145  
201 LEHLAKLVAVWSGSEVFRFSISGSGVFGGKLAIVVPP-----GVDPVQSTSLMQYPHVL 255  
146 IDVRLTLEPVLPLPDVRNPFHYNQNEPRMLVAMLYTFLPILTIGELTNSRFP-----VPIDELYTSPNE 255  
256 FDAQVDFPVIFSLDRLSTLYHLMPTD-TTSLVIMVYNDLINPYANDSNSSGCI- 310  
200 CRVLTRSPDPDFNYLVPPTLESKTKPFTLPILTIGELTNSRFP-----VPIDELYTSPNE 255  
311 -TVETKPGDPDFKFLHLPFG-----SMLTHGSPDLPKSSSLMIGNRYWSDIT 359  
256 SLVQOP 261  
360 DFVRP 365

RESULT 12  
VCWVFP9  
coat protein - feline calicivirus (strain F9)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 28-Jul-2000  
C:Accession: B43382; C45538; PQ0407; S23702  
R;Carter, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.  
Virology 190, 443-448, 1992  
A:Title: The complete nucleotide sequence of a feline calicivirus.  
A:Reference number: A43382; MUID:92410623; PMID:1529544  
A:Accession: B43382  
A:Molecule type: Genomic RNA  
A:Residues: 1-671 <CARL>  
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
R;Carter, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.

Arch. Virol. 122, 223-235, 1992  
A:Title: Identification and sequence determination of the capsid protein gene of feline  
A:Reference number: A45538; MUID:92117861; PMID:1731695  
A:Accession: C45538  
A:Molecule type: genomic RNA; protein  
A:Residues: 1-671 <CAR>  
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
A:Experimental source: strain F9  
A>Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)  
R;Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.  
J. Gen. Virol. 73, 2429-2433, 1992  
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel  
A:Reference number: PQ0407; MUID:93019069; PMID:1402818  
A:Accession: PQ0407  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>  
A:Cross-references: PIDN:AA23553.1; PID:g257083  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F;1177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.8%; Score 277.5; DB 1; Length 671;  
Best Local Similarity 31.4%; Pred. No. 1.6e-12;  
Matches 76; Conservative 38; Mismatches 97; Indels 31; Gaps 9;

30 EPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPGVGLLNLELGPENLPY 88  
145 EPSQMTAADMATGKSVDSWEAFPSFTSVNW-----STSETQGKILFKQSLGPLNLY 200  
89 LAHLRMYNGVAGGVEVQVLLAGNAFTAGKLVFAAAPPHPPLENISPGQIT---MPPHVI 145  
201 LEHLAKLVAVWSGSEVFRFSISGSGVFGGKLAIVVPP-----GVDPVQSTSLMQYPHVL 255  
146 IDVRLTLEPVLPLPDVRNPFHYNQNEPRMLVAMLYTFLPILTIGELTNSRFP-----VPIDELYTSPNESLV 259  
256 FDAQVDFPVIFSLDRLSTLYHLMPTD-TTSLVIMVYNDLINPYANDSNSSGCI- 313  
204 TRPSDPDPDFNYLVPPTLESKTKPFTLPILTIGELTNSRFP-----VPIDELYTSPNESLV 259  
314 TKPGDPDFKFLHLPFG-----SMLTHGSPDLPKSSSLMIGNRYWSDITDFVI 363  
260 QP 261  
364 RP 365

RESULT 13  
C48562  
coat protein - San Miguel sea lion virus (serotype 4)  
N:Alternate names: capsid protein  
C:Species: San Miguel sea lion virus  
C>Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
C:Accession: C48562  
R;Neill, J.D.  
Virus Res. 24, 211-222, 1992  
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel  
eius.  
A:Reference number: A48562; MUID:92410750; PMID:1529644  
A:Accession: C48562  
A:Molecule type: genomic RNA  
A:Residues: 1-703 <NEI>  
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAA16220.1; PID:g334888  
A>Note: sequence extracted from NCBI backbone (NCBIP:113567)  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F;89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.6%; Score 273.5; DB 1; Length 703;  
Best Local Similarity 28.8%; Pred. No. 3.3e-12;  
Matches 97; Conservative 51; Mismatches 150; Indels 39; Gaps 15;

8 AAPSDNGAAGLVPPEANNETM-----ALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAPNGE 63

```
Db 151 ASDGPGSABIVTEBOGTVVQOQAPAPATATLATAATGK-SVBOEWMTFFSYHTSNW 209
QY 64 FTVSPRNSPGEVLLNLELGPENLPYLHLRMYNGVAGVEVQVLLAGNAFTAGKLVFAA 123
Db 210 STV---ESQKILYSQALNPSINPYLDHIAKLYSTWSGGIDVRFTVSGSGVFGKLAALL 266
QY 124 VPPHP-PLENISPGQITMPEHVIDVRLPVLPLPDVRNFFHYNQOQNEPRMLVAML 182
Db 267 VPPGVEPIESVS---MLQYPHLVFDARQTEPFIPTIDIRKTLFHSMDTD-TTKLVINP 322
QY 183 YTPLRNSGGDDVFTVSCRVLRPSDFDNVLPPTLESKTKPFTLPILTIGELTNSRF 242
Db 323 Y-----ENGV-ENKTTCSITVETRPSADFTFALKPP--GSLIKHGSIP-----SDLIPRNS 371
QY 243 PVPIDELYTSNPNESLVQPOQ----NGRCALDGELOQ--TTQLLPTAICFSFRGRINQKVG 296
Db 372 AHWMGNRWSTISGFSVQVRVQSRNHRHFDSDTTTGMSTPYVPIEI-----KIQKVG 426
QY 297 EN---HVMNQVTNINGTPTDPTGDPAPLGTPTDFSG 330
Db 427 NKKWFHVIDTKALVGPID-DGWPDTTIPDETATNG 462
```

## RESULT 14

```
VCMWFF
coat protein - feline calicivirus (strain CFI/68 FIV)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text_change 20-Aug-1999
C:Accession: A40507; B40507; T09246
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.
J. Virol. 65, 5440-5447, 1991
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calicivirus
A:Reference number: A40507; MUID:91374597; PMID:1716692
A:Accession: A40507
A:Molecule type: Genomic RNA
A:Residues: 1-668 <NE1>
A:Cross-references: GB:M32819; NID:G323874; PIDN:AAA42925.1; PID:G323875
A:Accession: B40507
A:Molecule type: protein
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>
R:Neill, J.D.
submitted to the EMBL Data Library, April 1998
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.
A:Reference number: Z16626
A:Accession: T09246
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: Genomic RNA
A:Residues: 1-668 <NE3>
A:Cross-references: EMBL:U13992; NID:G3056875; PIDN:AAC13993.1; PID:G537256
A:Experimental source: strain CFI/68 FIV
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,439,459,461/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
Query Match 9.6%; Score 272.5; DB 1; Length 668;
Best Local Similarity 31.3%; Pred. No. 3.6e-12;
Matches 77; Conservative 37; Mismatches 93; Indels 39; Gaps 9;

QY 30 EPVAGASIAAPTGTQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPGEVLLNLELGPENLPY 88
Db 145 EPNAQMSTAADMATGKSDVSEWAEFFSHTSVNM-----STSETQKILFKQSLGILLNPY 200
QY 89 LAHLRMYNGVAGVEVQVLLAGNAFTAGKLVFAAVPPHPPLENISPGQIT---MPPHVI 145
Db 201 LTHLAKLYVWAGSGVDVRESISGSGVFGKLAIVVPP-----GIDPVQSTMLQYPHYL 255
QY 146 IDVRLTEPVLPLPDVRNFFHYNQOQNEPRMLVAMLYTPL-----RSNGSGDDVFTVS 199
Db 256 FPARQVEPFIPIDLRSTLYLHMSDTD-TTSLVMVNDLINPANDNSNGGCIV---- 310
QY 200 CRVLTRSPDPDFNVLPPTLESKTKPFTLPILTIGELTNSRFP-----VPIDELYTSNPN 255
```

```
Db 311 -TVETKPGDFPKHLKPPG-----SMLTHGSIPLSDLPKSSSLWIGNRPMWSDIT 359
QY 256 SLVVQP 261
Db 360 DVIIRP 365

RESULT 15
JQ2356
capsid protein - feline calicivirus (strain KCD)
C:Species: feline calicivirus
C>Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2356
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable regions
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2356
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:L09719; NID:G305107; PIDN:AAA16487.1; PID:G305108
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein
```

```
Query Match 9.4%; Score 266; DB 2; Length 668;
Best Local Similarity 29.6%; Pred. No. 1.1e-11;
Matches 79; Conservative 43; Mismatches 103; Indels 42; Gaps 11;

QY 11 SNDGAGLVPEANNET--MALEPVAGASIAAPLTQNNIIDPW-IRLNFVQAPNGEFTVS 67
Db 125 ADDGSI-VTPEQGTWVGVIAPNAQMTAADMATGKSDVSEWAEFFSHTSVNM-----S 179
QY 68 PRNSPGEVLLNLELGPENLPYLHLRMYNGVAGVEVQVLLAGNAFTAGKLVFAAVPPH 127
Db 180 TSETQKILFKQSLGILLNPYLHLAKLYVWAGSGIEVRFSISGSGVFGKLAIVVPP- 238
QY 128 PLENISPGQIT---MPPHVIDVRLTEPVLPLPDVRNFFHYNQOQNEPRMLVAMLYT 184
Db 239 ----GVDPVQSTMLQYPHLVFDARQVEPFIPIDLRSTLYLHMSDTD-TTSLVVMAYN 293
QY 185 PL-----RSNGSGDDVFTVSCRVLRPSDFDNVLPPTLESKTKPFTLPILTIGELT 238
Db 294 DLINPYANDNSNGGCIV-----TVETKPGSDFRHLKPPG-----SVLTHGSVP 338
QY 239 NSRFP-----VPIDELYTSNPNESLVQP 261
Db 339 SDLIPTSSLWIGNRYWTNITNFIIRP 365
```

Search completed: June 1, 2004, 13:55:28  
Job time : 11.323 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.4581 Seconds  
(without alignments)  
4985.230 Million cell updates/sec

Title: US-09-926-799-9  
Perfect score: 2863  
Sequence: 1 MKMASNDAPSDNGAASLVP.....VNQFYS LAPMTGTGNGRRRIQ 542

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1980	69.2	542	S60616	capsid protein - h
2	1833.5	64.0	539	S40111	capsid protein - h
3	1179	41.2	530	B37471	capsid protein - N
4	1159.5	40.5	546	B37491	major capsid prote
5	340	11.9	2344	S55399	genome polyprotein
6	339	11.8	2344	RRWRH	genome polyprotein
7	339	11.8	2344	S64740	genome polyprotein
8	326.5	11.4	576	A53982	capsid protein - E
9	290	10.1	668	JQ2354	capsid protein - f
10	279	9.7	668	VCWFF	coat protein - fel
11	276.5	9.7	671	VCWFF9	coat protein - fel
12	271.5	9.5	668	1 VCWFC	coat protein - fel
13	269.5	9.4	702	A48562	coat protein - San
14	268.5	9.4	668	JQ2356	capsid protein - f
15	240	8.4	703	C48562	coat protein - San
16	145.5	5.1	2333	1 GNNY2F	genome polyprotein
17	130	4.5	2332	1 GNNY4F	genome polyprotein
18	123.5	4.3	746	T46821	siderophore recept
19	123.5	4.3	746	A95420	RhtA Rhizobactin r
20	120.5	4.2	1742	S76110	hypothetical prote
21	119	4.2	701	A11501	internalin protein
22	118	4.1	2194	1 GNNY7F	genome polyprotein
23	116	4.1	545	S28117	gas-vesicle operon
24	115.5	4.0	3085	T00327	polyprotein - infe
25	114.5	4.0	785	S75819	general secretion
26	114	4.0	2336	S37077	genome polyprotein
27	111	3.9	940	2 D89723	protein F39D8.1b [
28	111	3.9	945	T21998	hypothetical prote
29	111	3.9	13055	2 T16580	hypothetical prote

RESULT 1

S60616  
capsid protein - human calicivirus (strain Melksham)  
C:Species: human calicivirus  
A:Variety: strain Melksham  
C:Date: 23-May-1997 #sequence\_revision 23-May-1997 #text\_change 28-Jul-2000  
C:Accession: S60616  
R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
A:Reference number: S60615; MUID:96136658; PMID:8533462  
A:Title: Capsid diversity in small round-structured viruses: molecular characterization  
A:Accession: S60616  
A:Molecule type: genomic RNA  
A:Residues: 1-542 <GRS>  
A:Cross-references: EMBL:X81879; NID:G976077; PIDN:CAA57462.1; PID:G976079  
A:Experimental source: strain Melksham  
A:Note: it is uncertain whether Met-1 or Met-3 is the initiator  
C:Superfamily: human calicivirus capsid protein  
C:Keywords: capsid protein; coat protein

Query Match	69.2%	Score	1980;	DB	2;	Length	542;
Best Local Similarity	67.6%	Pred. No.	1.2e-139;				
Matches	369;	Conservative	81;	Mismatches	88;	Indels	8;
Gaps	5;						
Qy	1	MMKASNDAPSDNGAASLVP	EGINMETPLEPVAGASIAAPVAGQTNIDPWRTNFVQAP	60			
Db	1	MMKASNDAPSDTGGAGLVPS	NEVMALEPVAGAAAPVTGTQTNIDPWIRANFVQAP	60			
Qy	61	NGEFTVSPRNGSGEILLNLEL	GPDLNPLYLAHLSRMVYAGGVVQVLLAGNAFTAGKTL	120			
Db	61	NGEFTVSPRNPAGEVLLNLEL	GPDLNPLYLAHLSRMVYAGGVVQVLLAGNAFTAGKLV	120			
Qy	121	FAAIPPNFLVDMISPAQITML	PHLIVDRTLEPTLPDVRNPFYHFNQPRMLVA	180			
Db	121	FAAIPPHFVENLSPKITMFP	PHVIVDRTLEPTLPDVRNPFYHFNQPRMIVA	180			
Qy	181	MLYPLRNGSGDDVFTVSCRV	LTPPTDFEFYLVPPSVESKTKPFTLPITISELTS	240			
Db	181	MLYPLRNGSGDDVFTVSCRV	LTPPTDFEFYLVPPSVESKTKPFTLPITISELTS	240			
Qy	241	RPPIEQLYTAPNETVNVQCON	GRCTLDGELQGTQTLSSAVCFLOGRVTADNGDNWQD	300			
Db	241	RPFPVIDQMTSPNEVISVQC	QNGRCTLGDELQGTQTLQVSGICAFKGEVTAHLHDNDHL	300			
Qy	301	NLLQUTYNGASDYDTDEVP	APLQDFSGMLYGVLTQDNVNVSTGEAKNAK---	GIYIS 357			
Db	301	NNVTITNLNGSPFDPSEDI	PAPLGVDFQGRVFGVISQDKQNAAGHSEPANRGHDVVP	360			
Qy	358	TTSGKFTPKIGSIGHLS-I	TEHVHFNQGSRTFPVGVADENTPFOQVLPVHAGSLANT	416			
Db	361	TYTAQYTKLGQIQIGTQWT	DDLTVNQPVKFTPVG--LNDTEHFNQMVVPRYAGALNLT	418			

QY 417 NLAPAVATPFGEOQLLFRSRVPCVQGLQGDADFICLLPQEWNNHVFQEAAPSQADVAL 476  
Db 419 NLAPSVAPVFFGERLLFRSHLP-LKGGVGNPA-IDCLLPQEWQHFQEAAPSMSSEVAL 476  
QY 477 IRYVNPDTGRTLFEAKLHRSFGITVSHGTCAVPLVVPNGHFRFDSWVNFQYSLAPMGTCN 536  
Db 477 VRYINPDTGRALFEAKLHRSFGITVSHGTCAVPLVVPNGHFRFDSWVNFQYSLAPMGTCN 536  
QY 537 GRRRQ 542  
Db 537 GRRRQ 542  
RESULT 2  
S40111  
capsid protein - human calicivirus (strain Bristol isolate B493)  
C:Species: human calicivirus  
A:Variety: strain Bristol isolate B493  
C:Date: 25-Dec-1994 #sequence\_revision 27-Feb-1997 #text\_change 28-Jul-2000  
C:Accession: S40111  
R:Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
submitted to the EMBL Data Library, December 1993  
A:Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de  
A:Reference number: S40111  
A:Accession: S40111  
A:Molecule type: genomic RNA  
A:Residues: 1-539 <GRE>  
A:Cross-references: EMBL:X76716; NID:G436410; PIDN:CAAS4134.1; PID:G436411  
A:Experimental source: human enteric calicivirus strain Bristol isolate B493  
C:Superfamily: human calicivirus capsid protein  
C:Keywords: capsid protein; coat protein  
Query Match 64.0%; Score 1833.5; DB 2; Length 539;  
Best Local Similarity 64.0%; Pred. No. 9.6e-129;  
Matches 349; Conservative 74; Mismatches 109; Indels 13; Gaps 6;  
QY 1 MKMASNDAAPSNDGAA--SLVPEGINETMPL--EPVAGASIAAPVAGQTNIIDPWIRTN 60  
Db 1 MKMASNDANPSDGAANLVPEVNEVNALEPVVGAALPAAPVAGQNVIDPWIRTNFVQAP 60  
QY 61 NGEFTVSPRNSPGBEILLNLELGPDLNPNLAHLSRMYNGYAGGVEQVLLAGNAFTAGKIL 120  
Db 61 GGEFTVSPRNPAGBILNLELGPDLNPNLAHLSRMYNGYAGGVEQVLLAGNAFTAGKVI 120  
QY 121 FAALPPNPLVDMISPAQITMLPHLIVDRTLEPMTPLDVRNVVYHFNQPPMRLLVA 180  
Db 121 FAAPVPPNPTGLSPSQVTMPFHTIIVDVRQLPEPLIPLDVRNNFYHYNQANDSTLKLIA 180  
QY 181 MLYTFLRNSGSDGDDVFTVSCRVLTPTDPDFEILVPPSVESKTKPFTLPILTIS 240  
Db 181 MLYTFLRANAGDDVFTVSCRVLTPTDPDFEILVPPSVESKTKPFTLPILTIS 240  
QY 241 RFPPIEOLYTAAPNETNVPVQNGRCTLDBLQGTQTLSSAVCFLOGRTVADGNDWDQ 300  
Db 241 RFPPIEOLYTGSSAFVVPVQNGRCTLDBLQGTQTLSSAVCFLOGRTVADGNDWDQ 300  
QY 301 NLLQTYNGASYPDTDEVPAPLGTQDFSGMLYGLVLTQDNVNVSTGEAKNAGIYITTS 360  
Db 299 YTMNLASQNWNSDPTSEIPAPLGTQDFSGMLYGLVLTQDNVNVSTGEAKNAGIYITTS 355  
QY 361 GKFTPKIGSGLHSITEH-VHPNQOSRFTPVGVAVD----ENTFPQWVLPHYAGSLALN 415  
Db 356 VHTFKLGSVQRTTNDNFQAGQNTKFTPVGVIQGDHQNKP-QWLLPNYSGRTHN 414  
QY 416 TNLAPAVATPFGEOQLLFRSRVPCVQGLQGDADFICLLPQEWNNHVFQEAAPSQADVA 475  
Db 415 VHLAPAVATPFGEOQLLFRSRVPCVQGLQGDADFICLLPQEWNNHVFQEAAPSQADVA 472  
QY 476 LIRYVNPDTGRTLFEAKLHRSFGITVSHGTCAVPLVVPNGHFRFDSWVNFQYSLAPMGTCN 535  
Db 473 LIRVFNPDGRLVFECKLHRSFGITVSHGTCAVPLVVPNGHFRFDSWVNFQYSLAPMGTCN 532  
QY 536 NGRRR 540

Db 533 TGERR 537  
RESULT 3  
B37471  
capsid protein - Norwalk virus  
C:Species: Norwalk virus  
C:Date: 24-Feb-1994 #sequence\_revision 25-Apr-1997 #text\_change 28-Jul-2000  
C:Accession: B37471  
R:Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.  
Virology 195, 51-61, 1993  
A:Title: Sequence and genomic organization of Norwalk virus.  
A:Reference number: A37471; MUID:93303939; PMID:8391187  
A:Accession: B37471  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: genomic RNA  
A:Residues: 1-530 <JIA>  
A:Cross-references: GB:M87661; NID:G1061311; PIDN:AABS0466.1; PID:G1061313  
A:Note: sequence extracted from NCBI backbone (NCBIP:134157)  
C:Superfamily: human calicivirus capsid protein  
Query Match 41.2%; Score 1179; DB 2; Length 530;  
Best Local Similarity 46.5%; Pred. No. 5.6e-80;  
Matches 256; Conservative 84; Mismatches 171; Indels 40; Gaps 12;  
QY 1 MKMASNDAAPSNDGAA--SLVPEGINETMPL--EPVAGASIAAPVAGQTNIIDPWIRTN 55  
Db 1 MMWASKDATSSVDGASGAGQLVPE-VNASDPLAMDVPVAGSSTAVATAGQVNPIDPWIRTN 59  
QY 56 FVQAPNGEFTVSPRNSPGBEILLNLELGPDLNPNLAHLSRMYNGYAGGVEQVLLAGNAFT 115  
Db 60 FVQAPNGEFTVSPRNSPGBEILLNLELGPDLNPNLAHLSRMYNGYAGGVEQVLLAGNAFT 119  
QY 116 AKGILFAAIIPNPLVDMISPAQITMLPHLIVDRTLEPMTPLDVRNVYHFNQPPMR 175  
Db 120 AKGILVSCIPGFGSHNLTAQATLPHVIAVDTLEPMTPLDVRNVYHFNQPPMR 179  
QY 176 MRLVAMLYTPLRNSGSDGDDVFTVSCRVLTPTDPDFEILVPPSVESKTKPFTLPILTIS 235  
Db 180 MRLVCMLYTPLRNSGSDGDDVFTVSCRVLTPTDPDFEILVPPSVESKTKPFTLPILTIS 239  
QY 236 ELTNSRFPPIEOLYTAAPNETNVPVQNGRCTLDBLQGTQTLSSAVCFLOGRTVADG 295  
Db 240 ELTNSRFPPIEOLYTAAPNETNVPVQNGRCTLDBLQGTQTLSSAVCFLOGRTVADG 296  
QY 296 DNWDQNLQLTYNGASYPDTDEVPAPLGTQDFSGMLYGLVLTQDNVNVSTGEAKNAGIY 355  
Db 297 DNWDQNLQLTYNGASYPDTDEVPAPLGTQDFSGMLYGLVLTQDNVNVSTGEAKNAGIY 344  
QY 356 ISSTGKFTPKIGSGLHSITEH-VHPNQOSRFT---PVGVAVDENTFPQWVLPHYAGS 411  
Db 345 VDTTPTDTPFVPHLGSIQANGIGSGNVGVLSWISPPSHPSGSGVD-----LWKIPNYGSS 398  
QY 412 LALNTNLAPAVATPFGEOQLLFRSRVPCVQGLQGDADFICLLPQEWNNHVFQEAAPS 470  
Db 399 ITEATHLADSVVPPGEGEVLVFFNSKMP-----GPGAYNLPCLLPQEVLSHASEQAPT 452  
QY 471 QADVALIRYVNPDTGRTLFEAKLHRSFGITVSHGTCAVPLVVPNGHFRFDSWVNFQY 528  
Db 453 VGEAALLHYVDPDTGRNLGEFKAYPDGFLTCVPNGASSGPGQQLPGLINGVVFVSWVSRFYQ 512  
QY 529 LAPMGTCNGRR 539  
Db 513 LKPVGTASSAR 523  
RESULT 4  
B37491  
major capsid protein [similarity] - Southampton virus  
N:Alternate names: orf2 protein  
C:Species: Southampton virus  
C:Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 18-Aug-2000



C:Accession: B37491  
R:Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
Science 259, 516-519, 1993  
A:Title: Sequence and genome organization of a human small round-structured (Norwalk-like  
A:Reference number: A37491; MUID:93142023; PMID:8380940  
A:Accession: B37491  
A>Status: not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-546 <LAM>  
A:Cross-references: GB:L07418; NID:g1236787; PIDN:AAA2984.1; PID:g295114  
A>Note: sequence extracted from NCBI backbone (NCBIP:123456)  
A:Note: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype  
C:Superfamily: human calicivirus capsid protein  
C:Keywords: glycoprotein  
F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 40.5%; Score 1159.5; DB 2; Length 546;  
Best Local Similarity 45.3%; Pred. No. 1.6e-78;  
Matches 258; Conservative 84; Mismatches 176; Indels 51; Gaps 17;

QY 1 MMWASNDAAAPSNDGAA---SLVPEGIN--ETWPLEPVGASTAAPVAGQTNIIDPWRTN 55  
DB 1 MMWASKDAPQASDAGSAGQVPE-VNTADPLPMEPVAGPTTAVATAGQVNMIDPWVNN 59  
QY 56 FVQANGGETVSPRNSPGBEILLNLDELGPDLNPLHLASRMVNGYAGGVEQVLLAGNAPT 115  
DB 60 FVQSQGEPTISPNNTPGDILFDLQGLHNLPLSHLSQMYNGWGMVRILLAGNAPS 119  
QY 116 ACKILFAAIPNPLVDMLSPAQITMLPHLIVDVRTLEPMTPLPDVNVVYHFNNOPOR 175  
DB 120 AKIIVCCVPGPFTSSSLTAQATLPHVIAVDRLEPIEMPLEDRVNLVHTNQN-QPT 178  
QY 176 MELVAMLYPLRS-NGSGD-DVFTVSCRVLTPDFFFIYLVPPSVESKTKPFTLPILT 233  
DB 179 MELVCMLYPLRTGGSGNSDSFVAGRVLTAPSSDFSFLVLPPTISQKTAFTVPNIP 238  
QY 234 ISELNRPPIPIEQLYTAPNETNVVQNGRCTLDELQGTQLLSAVCFLOQRTVAD 293  
DB 239 LQTLNSRPPSIIQGMILSPDASQVVFQNGRLDGLLGTTPATSGOLFVRGKI--- 295  
QY 294 NGDNWDNLQLYTNGASYDTPDEVPALGQDFSGMLYGLVLTQDNVNVSTGEAKNAG 353  
DB 296 ---NQARTLNITEVDGKPFMAFDS-PAPVGPDPFGKCDWHMRISKTPN-NTGSGDPMRS 350  
QY 354 IYISTSGKFTPKIGSIGLSITEHVHP-----NQSRFTTPVGVAVDENTPFOQ 402  
DB 351 VSVQTNVQGFVPHLSIQDFEVEN-HPTGDIYGTIEWISQFS--TPPG-----TDINL 400  
QY 403 WVLPHYAGSLALNTNLAPAVAPTFPGEQLLFRRSRVPCVQGLQGQDA--FIDCLLPQEWV 460  
DB 401 WEIPDYGSSLSQAANLAPVPVFPFGFGEALVYFVSAPP---GPNRSAPNDVPCLLPQEI 457  
QY 461 NHFYQEAAPSQADVALIRVYVNDTGTLEFAKLHRSGLFTV--SHTGAYPLVVPNGHPR 518  
DB 458 THFVSEQAPTMDAALLHYVDPTNRNLGEFKLYPGGYLTCVPGVGVAGPQOPLNGVFL 517  
QY 519 FDSWVNOFSLAPMTGN-----GRRRI 541  
DB 518 FVSWVSRFYQLAPVGTASTARGELGVRI 546

RESULT 5  
S55399  
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)  
C:Species: rabbit hemorrhagic disease virus  
A:Variety: isolate BS89  
C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 18-Jun-1999  
C:Accession: S55399  
R:Rossi, C.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: S55399  
A:Accession: S55399  
A:Molecule type: genomic RNA

A:Residues: 1-2344 <ROS>  
C:Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641  
A:Experimental source: isolate BS89  
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C:Keywords: polyprotein

Query Match 11.9%; Score 340; DB 2; Length 2344;  
Best Local Similarity 26.4%; Pred. No. 1.7e-16;  
Matches 141; Conservative 70; Mismatches 209; Indels 114; Gaps 22;

QY 19 VPEGINEMPLEPVGASIAAPVAG-QTNIIDPMI----- 52  
DB 1765 VMEGKARTAPQGEAACTATTASVPTTGDGLDPGVVATTSVVTAESSASIAIATAGIGGPP 1824  
QY 53 -----RTNFVQAPNGEFTVSPRNSPGBEILLNLDELGPDLNPLHLASRMVNGYAGGV 103  
DB 1825 QVDDQOQETWRTNFY--NDVFTWSVADAPGSILYTVQHSPPQNNPFTAVLSQMYAGWAGGM 1882  
QY 104 EVOVLLAGNATAGKILFAAIPNPLVDMLSPAQITMLPHLIVDVRTLEPMTPLPDVNR 163  
DB 1883 QFRFTVAGSGVFGGRLVAIVPPG--IEIGPLEVRQPFHVVIDARSLEPVTITPMDLRP 1940  
QY 164 VFYHENNOPQPMRLVAMLYPLRSNGSGDDVFTVSCRVLTPDFFFIYLVPPSVESK 223  
DB 1941 NMYHTGDPGLVPTLVSYNNL-INPFGSSTSAIQVTETRPSEDFEFVMTIRAFS--SK 1997  
QY 224 TKPFTLP--ILTISELT-----NSRPPPIEQLYTAPNETNVVQNGRCTLDELQGTQ 277  
DB 1998 TVDSISPAGLLTPVLTGVGNDNRWNGQIVGLQVPEGFST--C-NRHNWLNSTGVWSS 2054  
QY 278 LSSAVCFLOQRTVADNGNDQNLQLYTNGASY--DPTDEVPAPIGTQD-----FSG 330  
DB 2055 PRFADIDHRRG-SASYPGSN-ATNVLQFWYANAGSAVDNPISQV-APDGFDPMSFVPFG 2111  
QY 331 -----MLXGVLTQDN---VNVSTGRKN-----AKGIYI 356  
DB 2112 PCIPAAGWVGCAIWNSSGAPNVTVQAYELGFATGAPGNLQPTNTSGAOTVAKSIIYA 2171  
QY 357 STT-----SGKFTPKIGSIGLSITEHVHPNQSRF-----TPVGVAVDENTPFOQWYL 405  
DB 2172 VVTGTAQNPAGLFVWASGVISTPNANAITYPQDRIVTTPGTPTAAAPVGNKTPIM--- 2227  
QY 406 PHYAGSLALNTNLAPAVAPTFPGEQLLFRRSRVPCVQGLQGQDAFIDCLLPQEW 459  
DB 2228 ---FASVVRRTGDVNVATAGSANGTQGTGSOPLPTVITGL-SLNYSYSSALMPQOF 2277

RESULT 6  
RRWRH  
genome polyprotein - rabbit hemorrhagic disease virus  
N:Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C:Species: rabbit hemorrhagic disease virus  
C:Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998  
C:Accession: A41039  
R:Meyers, G.; Wirblich, C.; Thiel, H.J.  
Virology 184, 664-676, 1991  
A:Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing  
A:Reference number: A41039; MUID:91361557; PMID:1840711  
A:Accession: A41039  
A:Molecule type: genomic RNA  
A:Residues: 1-2344 <MEY>  
C:Cross-references: GB:M67473  
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C:Keywords: nucleotidyltransferase

Query Match 11.8%; Score 339; DB 1; Length 2344;  
Best Local Similarity 27.5%; Pred. No. 2e-16;  
Matches 145; Conservative 71; Mismatches 214; Indels 98; Gaps 25;

QY 8 AAPSNDGASL-----VPEGINEMPLEPVA-----GASIA-APVAGQTNIID--- 49  
DB 1772 AAPQGEAACTATTASVPTTGDMDPGVVATTSVVTAESSASIAIATAGIGGPPQVDOOE 1831

50 PWIRNFVQANGGEFTVSPRNSPGEILLNLLELGLPDLNPLYLALHLSRMVNGYAGGVQVLL 109  
 1832 TW-RNFYY--NDVFTWSVADAPGILTYVQHSQNNPFTAVLSQMYAGWAGGQRFV 1888  
 110 AGNAFTAGKILFAAIPPNFLVDMISPAQITMLPHLIVDVRLTLEIMPLPDRNVFVFN 169  
 1889 AGSGVFGGLVRAVIPP--IEIGPGLVRFQPHVVIDARSLPFTITMPDLRPNMYHT 1946  
 170 NOPQPRMLVAMLYTPLSRSGDDVFTVSCRVLTRPTDPDEFIYLVPPSVESKPKFTL 229  
 1947 GDFGLVPTLVLSVYNL--INPFGSTSAIQVTVETRPSEDFEVMIRAPS--SKTVDSIS 2003  
 230 P--ILITISELT---NSRFPPIEQLYTAPNETNVVQCNGRCTLDGELGTQTOLLSSAV 283  
 2004 FAGLITPTLVLTGVGNDNRWNGQIVGLQPVPGFEST--C-NRHWNLNGSTYGWSSPRFGDI 2060  
 284 CFLQRTVADNGDNDQNLQLYTPNGASY--DPTDEVPAPLGTQD-----FSG----- 330  
 2061 DHRRG--SASYSGSN--ATNVLFQWYANAGSAIDNPISQV--APDGFPMDFVFPNGGIPAA 2117  
 331 --MLYGLVLTQDN---VNVSTGEAKN-----AKGIYISTT--- 359  
 2118 GWVFGGAIWNSGAPNVTVQAYELGFATGAPGNLQPTNTSGAQTVAKSIYAVVTGTA 2177  
 360 ---SGKFTPKIGSIGLHSITEHVHPNQQRFP-----TPGVAVDENTFPQWVLPHYAGS 411  
 2178 QNPAGLFVWASGIISTENAGRIITYTPQDRIVTTPGTAAAPVGNKTPIM-----FAS 2230  
 412 LALNTNLAPAVATFPGEQLLFRSRVPCVQGLQGDADIDCLLPQEW 459  
 2231 VVRRTGDNVATAGSANGTQYGTGSQLPVTIGL--SLNNYSALMPGQF 2277

RESULT 7  
 S64740 genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)  
 N;Contains: VP60 protein  
 C;Species: rabbit hemorrhagic disease virus  
 A;Variety: isolate AST/89  
 C;Date: 12-Jul-1996 #sequence revision 26-Jul-1996 #text\_change 20-Jun-2000  
 C;Accession: S64740; S46944; S49018; S65012  
 R;Boga, J.; Martin-Alonso, J.; Boga, J.; Parra, F.  
 submitted to the EMBL Data Library, May 1995  
 A;Description: Genomic organization of rabbit hemorrhagic disease virus determined by di  
 A;Reference number: S64740  
 A;Accession: S64740  
 A;Molecule type: genomic RNA  
 A;Residues: 1-2344 <CAS>  
 A;Cross-references: EMBL:Z49271; NID:gl182032; PIDN:CAA89265.1; PID:gl150552  
 A;Experimental source: isolate AST/89  
 R;Boga, J.; Casais, R.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.  
 submitted to the EMBL Data Library, July 1993  
 A;Description: Molecular cloning, sequence and expression of the capsid protein gene f  
 A;Reference number: S46944  
 A;Accession: S46944  
 A;Molecule type: genomic RNA  
 A;Residues: 1650-2344 <BOG>  
 A;Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAAB08081.1; PID:g515623  
 A;Experimental source: isolate AST/89  
 R;Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.  
 Virus Res. 27, 219-228, 1993  
 A;Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp  
 A;Reference number: S49018; MUID:93255896; PMID:8488721  
 A;Accession: S49018  
 A;Molecule type: genomic RNA  
 A;Residues: 1650-1796 <PAR>  
 A;Cross-references: EMBL:Z24757  
 A;Experimental source: isolate AST/89  
 A;Accession: S65012  
 A;Molecule type: protein  
 A;Residues: 1767-1779;1875-1877,'X',1879-1881,1936-1938,'X',1940-1941 <PAM>  
 C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
 C;Keywords: blocked amino end; polyprotein

Query Match 11.8%; Score 339; DB 2; Length 2344;  
 Best Local Similarity 26.9%; Pred. No. 2e-16;  
 Matches 144; Conservative 61; Mismatches 213; Indels 118; Gaps 20;

19 VPEGINETMPLFPVAGASTAAPVAG-QTNIIIDPMI----- 52  
 1765 VMEGKARTAPOGEAAGTATTASVPGTTTGDMDGPGVVAITVVVTAENSSASATATAGIGOPP 1824  
 53 -----RTNFWQAPNGEFTVSPRNSPGETILLNLELGLPDLNPLYLAHLSRMYNGYAGGV 103  
 1825 QQVDQETWRITFY--NDVFTWSVADAPGSLTYTVQHSQNNPFTAVLSQMYAGWAGGM 1882  
 104 EVQVLLAGNAFTAGKILFAAIPPNFLVDMISPAQITMLPHLIVDVRLTLEIMPLPDRNV 163  
 1883 QRFIVVAGIGVFGGLVRAVIPP--IEIGPGLVRFQPHVVIDARSLPFTITMPDLR 1940  
 164 VFYHFNQPPQPRMLVAMLYTPLSRSGDDVFTVSCRVLTRPTDPDEFIYLVPPSVESK 223  
 1941 NMVHTGDPGLVPTLVLSVYNL--INPFGSTSAIQVTVETRPSEDFEVMIRAPS--SK 1997  
 224 TKPFTLP--ILITISELT---NSRFPPIEQLYTAPNETNVVQCNGRCTLDGELGTQT 277  
 1998 TVDSISFAGLITPTLVLTGVGNDNRWNGQIVGLQPVPGFEST-----CNRHWNLNGSTY 2050  
 278 LLSAVCFLOQ--RTVADNGDNDQNLQLYTPNGASY--DPTDEVPAPLGTQD-----F 328  
 2051 GWSPPREGDIGHRRGSASYPGNATNVLFQWYANAGSAIDNPISQV--APDGFPMDFVFP 2109  
 329 SG-----MLYGLVLTQDN---VNVST-----GEAKNAKGI 354  
 2110 NGPPIAAGVWFGAIWNSGAPNVTVQAYELGFATGAPGNLQPTNTSGAQTVAKSI 2169  
 355 YISTT-----SGKFTPKIGSIGLHSITEHVHPNQQRFP-----TPGVAVDENTFPQW 403  
 2170 YAVVTGTAQNPAGLFVWASGVISTPSANAITYPQDRIVTTPGTAAAPVGNKTPIM-- 2227  
 404 VLPHYAGSLALNTNLAPAVATFPGEQLLFRSRVPCVQGLQGDADIDCLLPQEW 459  
 2228 ----FASVVRRTGDNVATAGSANGTQYGTGSQLPVTIGL--SLNNYSALMPGQF 2277

RESULT 8  
 A53982 capsid protein - European brown hare syndrome virus  
 C;Species: European brown hare syndrome virus  
 C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 28-Jul-2000  
 C;Accession: A53982  
 R;Wirdlich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Eskens, U.; Haas, B.; Thiel, H.  
 J. Virol. 68, 5164-5173, 1994  
 A;Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease  
 A;Reference number: A53982; MUID:94309183; PMID:7518531  
 A;Accession: A53982  
 A;Status: preliminary  
 A;Molecule type: mRNA  
 A;Residues: 1-576 <NIR>  
 A;Cross-references: GB:U09199  
 C;Superfamily: human calicivirus capsid protein

Query Match 11.4%; Score 326.5; DB 2; Length 576;  
 Best Local Similarity 26.9%; Pred. No. 2.1e-16;  
 Matches 127; Conservative 67; Mismatches 157; Indels 121; Gaps 24;

3 MASNDAAPSNDGAASLVPEGINETMPLFPVAGASTAAPVAGQTNIIIDPWIRTNFVQAPNG 62  
 31 VASTDVVTADNVAAS-----VATAGTGGP-PQASQESW-RVNFY--ND 72  
 63 EFTVSPRNSPGETILLNLELGLPDLNPLYLAHLSRMYNGYAGGVQVLLAGNAFTAKILPA 122  
 73 VFTWSVTDAPGSLTYSVQHSQNNPFTQVLSQMYAGWAGGMQRFRTVAGSGIFGGLVCA 132  
 123 AIPPNFLVDMISPA-QITMLPHLIVDVRLTLEIMPLPDRNVVVFVHFNQPPQPRMLVAM 181

Db 133 IIPPGI---QIQGLEVRQPHVVIDARSLPVTITMPDLRPMVHTPGDGLVPLTVS 189  
QY 182 LYTPLASNGDDVFTVSCRVLTRPTDFEFYLYPPSVESKT---KP---FTLPILTI 234  
Db 190 VYNNL-INFPGGTSAIQVTVETRPSEDEFVLIRAPS--SKTVDSVNSWLLTTPVLT- 245  
QY 235 SELTNSRFPPIEQLYATANEVNVQCNGRCTLDGELQ----- 274  
Db 246 GAGSDNRWGAIPVGLQPVPGFST---SNRHNMNGSTYGWSRPFDDIDHPSGNVSYPT 302  
QY 275 -----TTQLSS-----AVCFLLQGRVADN-----GDNDQD- 301  
Db 303 GSATNTIETWYANAGTATTNPISNTAPDGFPMGAIPP-SGTITPGAWVGFQVWNASN 361  
QY 302 -----LLOLYTPNGA--SYDPTDEVPAPLGTQDFSMGLYGLVLTQDNVNVSTGEAKN 350  
Db 362 GTPYVGTQVAYELGFANGAPSSIRP---VTTTGAQLVAKSIYGVIAIQNS-SAGIIFL 417  
QY 351 AKGIYIST---TSGKFTPKIGSLHSITEHVHPNQOSRFTPVGVAVDENTP 399  
Db 418 SKGM-VSTPGVAATTVPQSAI---VT-----TPGTPVAAPIGKNTP 456

RESULT 9  
JQ2354  
capsid protein - feline calicivirus (strain NADC)  
C:Species: feline calicivirus  
C>Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
C:Accession: JQ2354  
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.  
J. Gen. Virol. 74, 2519-2524, 1993  
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable  
A:Reference number: JQ2354; MUID:94065683; PMID:7504075  
A:Accession: JQ2354  
A:Molecule type: mRNA  
A:Residues: 1-668 <SEA>  
A:Cross-references: GB:L09718; NID:G305104; PIDN:AAA16485.1; PID:G305105  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein

Query Match 10.1%; Score 290; DB 2; Length 668;  
Best Local Similarity 23.4%; Pred. No. 1.4e-13;  
Matches 138; Conservative 88; Mismatches 201; Indels 164; Gaps 30;

QY 11 SNDGAASLVPEG-INETMLEPVAGASIAAPVAGQNIIDPW-IRTNFVOAPNGEFTVSP 68  
Db 125 TDDGSITILEQGTVGVIAEPAQMSAQMSTADMATKGSVDSEWEAPFSTSVNW----ST 180

QY 69 RNSPGEILNLELGPDLNPLYLAHLSRMVNGYAGGVEVQVLLAGNAFTACKILFAAIPPNF 128  
Db 181 SETQKILFKQSLGPLLNPYLEHLSKLYVAVSGSVSEVRFSGSGVFGGKLAIVVPPG- 239

QY 129 LVDMISPAQITMLPHLVDVRLTLEPMTLPDVRNVFVFNQPO-PRMRLVAMLYTLP- 186  
Db 240 -VDPVQSTMLQPHVLPDARQVDFVIFSDPLRSTLYHL--MPDPTTSLVIMVYNDLI 296

QY 187 -----RNSGSGDDVFTVSCRVLTRPTDFEFYLYVPP-----SVESK----- 223  
Db 297 NPYANDSNSGGIV-----TVETKPGDPFKFLLPKPGSMLTGSPVSLIPKSSSLWIG 351

QY 224 -----TKPFTL-----PI-LTISELTNSRFPPIEQY 250  
Db 352 NRYWSDITDFVVRFPFQANRHFNFQETAGWSAPRFRPITITISESKSGKLIGVATDY 411

QY 251 TAPN-----ETNVQCNGRCTLDGELQGTQLLSSAVCFLOGRTVADNGNDONLLO 304  
Db 412 IVFPIPDGWPDTTIAE-----DLTPAGDYAIT-----SGNGD----- 444

QY 305 LTVPNGASDPTDEVPAPLGTQDFSGM-LYGVLTQDNVNVSTGEAKNAKGIYSTT---S 360  
Db 445 ---ITGSEYDSTEVK---NNTNFRGMVYCGSLQR-----ANGDKKISNTAFITTAKEG 494

QY 361 GKFTPK-----IGSIGLHSITEHVHPNQOSR-----FTPVG---VAVDENTPFQOWYL 405

Db 495 NKIRPENTIDMTKIAVYQDT-HVGBEVQTSDDALLALGYTGIEQAIGSDRDRVVRISVL 553  
QY 406 PHYAGSALANTNLAPAVAPTFPGEQLLPPRSRVPVCVQGLQGQDAFIDCULL---PQSWVNH 462  
Db 554 PEVGAR-----CGNHPIFYKNSIKLGYVIRSIDVFNSQLHTSRQLSLNH 598  
QY 463 FYOEAPASQADVALLRYNPDGRTLFEAKLHRS--FITVSHTG--AYPL 509  
Db 599 YLL-----SPDSPAVYRII--DSNGSWFQIDSEGFSFVGISIGKLEYPL 643

RESULT 10  
VCWVFF  
coat protein - feline calicivirus (strain CFI/68 FIV)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: A40507; B40507; T09246  
R:Neill, J.D.; Reardon, I.M.; Heinrichson, R.L.  
J. Virol. 65, 5440-5447, 1991  
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calici  
A:Reference number: A40507; MUID:91374597; PMID:1716692  
A:Accession: A40507  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <NEI>  
A:Cross-references: GB:M32819; NID:G323874; PIDN:AAA42925.1; PID:G323875  
A:Accession: B40507  
A:Molecule type: protein  
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>  
R:Neill, J.D.  
submitted to the EMBL Data Library, April 1998  
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.  
A:Reference number: Z16626  
A:Accession: T09246  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <NE3>  
A:Cross-references: EMBL:U13992; NID:G3056875; PIDN:AAC13993.1; PID:G537256  
A:Experimental source: strain CFI/68 FIV  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.7%; Score 279; DB 1; Length 668;  
Best Local Similarity 23.2%; Pred. No. 9.2e-13;  
Matches 137; Conservative 81; Mismatches 210; Indels 162; Gaps 28;

QY 11 SNDGAASLVPEGINETMP-----LEPVAGASIAAPVAGQNIIDPW-IRTNFVOAPNGEFT 65  
Db 125 ADDGSITTPEQG---TMVGVIAEPNAQMSAQMSTADMATKGSVDSEWEAPFSTSVNW--- 178

QY 66 VSPRNSPGEILNLELGPDLNPLYLAHLSRMVNGYAGGVEVQVLLAGNAFTACKILFAAIP 125  
Db 179 -STSETQKILFKQSLGPLLNPYLEHLSKLYVAVSGSVSDVRFSGSGVFGGKLAIVVPP 237

QY 126 PNFLVDMLSPAQITMLPHLVDVRLTLEPMTLPDVRNVFVFNQPO-PRMRLVAMLYTLP 185  
Db 238 PG-IDPVQSTMLQPHVLPDARQVDFVIFSDPLRSTLYHLMSDTD-TTSLVIMVYND 294

QY 186 L-----RNSGSGDDVFTVSCRVLTRPTDFEFYLYVPP-----SVESK----- 223  
Db 295 LINPVANDSNSGGIV-----TVETKPGDPFKFLLKPPGSMLTGHSIPESDLIPKSSSLW 349

QY 224 -----TKPFTL-----PI-LTISELTNSRFPPIEQ 248  
Db 350 IGNRFWSDITDFVVRFPFQANRHFNFQETAGWSAPRFRPITITISESKSGKLIGVAT 409

QY 249 LYTAPE-----ETNVQCNGRCTLDGELQGTQLLSSAVCFLOGRTVADNGNDONL 302  
Db 410 DVIIVGIPDGWPDT-----TIPGELVPV-----GDYAITNGTND--- 444

QY 303 LQLYTPNGASDPTDEVPAPLGTQDFSGM-LYGVLTQ-----DNNVNVSTGEAKNAK 353

```
Db 445 ----ITTAQYDAAEI---RNNTFRCGMYICGSLQRAWGDKKISNTAFITTTGVDGAXL 497
QY 354 IYISTSGKFTPKKISIGLSHTEHVHNQOS-----RTPVG---VAVDENTPQQWVLP 406
Db 498 IPSNTID---QTKIAVFQDTHANKHVQSDDTLALLGYTGIGEEAIGHADRVRVRSVLP 554
QY 407 HYAGSLALNTNLAPAVATPFCEQLLFFRSRVPVCVQGLQGDQDAFIDCLL---PQSVNHF 463
Db 555 ERGAR-----GGNHPIPHKNSIKLGVVIRSIDVFNSQLHRSQILHSLNHY 599
QY 464 YQEAAPSQADVALIRVNPDTORTLFEAKLHRSG--FTVSHTG--AYPL 509
Db 600 LL-----SPDSFAVYRII--DSNGSWFDIGIDNDGFSFVGSSIGKLEPFL 643
```

## RESULT 11

```
VCWMF9
coat protein - feline calicivirus (strain F9)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text_change 28-Jul-2000
C:Accession: B43382; C45538; FQ0407; S23702
R:Cartier, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.
Virology 190, 443-448, 1992
A:Title: The complete nucleotide sequence of a feline calicivirus.
A:Reference number: A43382; MUID:92410623; PMID:1529544
A:Accession: B43382
A:Molecule type: genomic RNA
A:Residues: 1-671 <CAR1>
A:Cross-references: GB:M86379; NID:G323877; PIDN:AAAY9327.1; PID:G323879
R:Cartier, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.
Arch. Virol. 122, 223-235, 1992
A:Title: Identification and sequence determination of the capsid protein gene of feline
A:Reference number: A45538; MUID:92117861; PMID:1731695
A:Accession: C45538
A:Molecule type: genomic RNA; protein
A:Residues: 1-671 <CAR2>
A:Cross-references: GB:M86379; NID:G323877; PIDN:AAAY9327.1; PID:G323879
A:Experimental source: strain F9
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBI:77462)
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.
J. Gen. Virol. 73, 2429-2433, 1992
A:Title: The cloning, sequencing and expression of a major antigenic region from the fel
A:Reference number: FQ0407; MUID:93019063; PMID:1402818
A:Accession: FQ0407
A:Status: nucleic acid sequence not shown
A:Molecule type: mRNA
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>
A:Cross-references: PIDN:AA823553.1; PID:G257083
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
Query Match 9.7%; Score 276.5; DB 1; Length 671;
Best Local Similarity 25.1%; Pred. No. 1.4e-12;
Matches 131; Conservative 68; Mismatches 171; Indels 151; Gaps 28;
```

```
QY 11 SNDGAASLVPEGINETMP-----LEPVAGASTAAPVAGQTNIIDPW-IRTNFVQAPNGEFT 65
Db 125 ADDGSITAEQ-----TWVGGVIAEPSQAQMTAADMATGKSDVSEWAEFFSFHTSVNW--- 178
QY 66 VSPRNSPGEILLNLELGLDPLNLYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKILFAAIP 125
Db 179 -STSTQKILPKQSLGPNLNYLEHLAKLYVAMSGSLEVRSISGSGVFGKLAIVP 237
QY 126 PNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFHFNNQOPQPRMRLVAMLYTP 185
Db 238 FG--VDPVQSTMLQYPHVLPDARQVPIFCLPDLRSTLXHLMSDITD--TTSLVIMVYND 294
QY 186 LRSNGSGDDVFTVSC--RVLTPTPDPEFYLVP-----SVES-----TK----- 224
Db 295 L-INPYANDANSGCCIVTVETKPGDFKPHLLKPGKMLTHGSIPLSDILPKTSSILWGNR 353
```

```
QY 225 -----KPFTL-----PI-LTISELTNSRFFPIPIEQLYTA 252
Db 354 YMSDITDFVIRPFVQANHFDPNQETAGWSTPRFRPISVITEQNGAKLIGVATDIY 413
QY 253 PN-----ETNVVQCGNGRCTLDGEL-----QGTOLLSSAVCFLOGETVAD--N 294
Db 414 PGIPDGPDT-----TIPGELIPAGDVAITNGTGNDITAT-----GYDTADIKN 459
QY 295 GDN-----W-DONLLQLLTPNGASVD--PTDEVPAPLGTQDFSGMLYGVLTQD 339
Db 460 NTNFRGMYICGSLQRAWGDKKISNTAFITATLDGNNKINPCNTIDOSKI---VVFQD 516
QY 340 N-----VNVS-----TGEAKNAKG-----IYISTSGKFTPKKISIGLSHTEHV 379
Db 517 NHVGKKAQTSDDTLALLGYTGIGEAIGSDRDRVVRIST-----LPETGARGN----- 565
QY 380 HPNQQRFTPVGVAVDENTPQQWVLPVHYAGSLALNTNLAP 420
Db 566 HPFYKNSIKLGVVIRSIDVFNSQLHRSQILHSLNHYLLP 605
```

## RESULT 12

```
VCWMF9
coat protein - feline calicivirus (strain Japanese F4)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C>Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C:Accession: B40481
R:Tohya, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamaza
Virology 183, 810-814, 1991
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.
A:Reference number: A40481; MUID:91306470; PMID:1853578
A:Accession: B40481
A:Molecule type: genomic RNA
A:Residues: 1-668 <TOH>
A:Cross-references: GB:D90357; NID:G221264; PIDN:BAA14371.1; PID:G221266
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted
```

```
Query Match 9.5%; Score 271.5; DB 1; Length 668;
Best Local Similarity 22.5%; Pred. No. 3.3e-12;
Matches 130; Conservative 77; Mismatches 193; Indels 179; Gaps 25;
```

```
QY 30 EPVAGASTAAPVAGQTNIIDPW-IRTNFVQAPNGEFTVSPRNSPGEILLNLELGLDPLNPY 88
Db 145 EPSAQMTAADMASGKSDVSEWAEFFSFHTSVNW---STSETQKILFKQSLGPNLNPY 200
QY 89 LAHLSRMVNGYAGGVEQVLLAGNAFTAGKILFAAIPDPLNLYDMISPAQITMLPHLIVDV 148
Db 201 LEHLAKLYVAMSGSLEVRSISGSGVFGKLAIVVPPG--VDPVQSTMLQYPHVLPFA 258
QY 149 RTLEPIMTLPDVRNVFHFNNQOPQPRMRLVAMLYTP-----RNSGSGDDVFTVSCR 202
Db 259 RQVEPVIFIPDLRSTLXHLMSDITD--TTSLVIMVYNDLINPYANDNSGCCIV-----TV 312
QY 203 LTRPTPDPEFYLVP-----SVESK-----TKPFTL----- 229
Db 313 ETKPGDFKPHLLKPGKMLTHGSIPLSDILPKSSSLWNTGNRWITDITDFVIRPFVQANR 372
QY 230 -----PI-LTISELTNSRFFPIPIEQLYTAFTNPNVTVQCGNGRCTLGDE 271
Db 373 HFDNQETAGWSTPRFRPITITISEKNGSKLIGVATDIYIPGIP----- 417
QY 272 LQGTOLLSSAVCFLOGETVANGDNW--DONLLQLLTPNGASVDPTDEVPAPLGTQDFSG 330
Db 418 -----DGWPDPTTIADKLIPAGD----- 434
QY 331 MLYGVLTQDNVNVSTGEAKNAKGIYISTSGKFTPKIGSI-----GLHST----- 376
Db 435 --YSITTTGENDIKTAQADYDAAVKNQNTNFRGMYICGSLQRAWGDKKISNTAFITAIR 492
```

QY 377 --EHVHPNQSRFTPVGVADVENTPQQWV-----LPHYA--GSLAINTNLAPAV-- 422  
Db 493 DGNIEKPSNTIDMTK--LAVYQDTHVEQSVQTSDDTLALLGTYTGIGEEAIGNSRDVVRI 550  
QY 423 -----APTTPGQQLFFRSRVCVQGLQODAFIDCLL---PQWVNHFYQEAAPSQADV 474  
Db 551 SVLPAGARGGNHPIFYKNSIKLVIRSIDVFNQSLHTSRLSLNHY---LLPPDS-F 606  
QY 475 ALIRVYNPDGTGTLFEAKLHRSQ--FITVSHTG--AYPL 509  
Db 607 AVYRII--DSNGSWFDIGDSGFSFGVSDIGKLEFPL 643

RESULT 13

A48562  
Query Match 9.4%; Score 269.5; DB 1; Length 702;  
Best Local Similarity 25.0%; Pred. No. 5e-12; Mismatches 169; Indels 97; Gaps 21;  
Matches 108; Conservative 58;  
N:Alternate names: capsid protein  
C:Species: San Miguel sea lion virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
C:Accession: A48562  
R:Neill, J.D.  
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lions.  
A:Reference number: A48562; MUID:92410750; PMID:1529644  
A:Accession: A48562  
A:Molecule type: genomic RNA  
A:Residues: 1-702 <NEI>  
A:Cross-references: GB:M87481; NID:q334882; PIDN:AAAL6217.1; PID:q334884  
A:Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBI:P:113565)  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

QY 8 AAPSNDGAASLVPEGINMTPLPVPAGAS----IAAPVAGQTNIIID-PWIRTNFVQ---A 59  
Db 151 AESDGPAGADIVTEQGTVVQOQPVPAQSAITLAAASTGKT--VDCEN--TTFFSYHTA 206  
QY 60 PNGETVSPRNSPGBIILNLELGPDLNLYLAHLSRMVYAGGVQVLLAGNAFTAGKI 119  
Db 207 VNWSTT---EAQGKILFSLRSLNLYLAHLSRMVYAGGVQVLLAGNAFTAGKI 119  
QY 120 LFAAIPPNFLVDMISPAQITMLPHLIVDVTLEPIMTPLPDVNRVYHFNQPPQPRMLV 179  
Db 263 AALIVPPG--IEPVESPTMLQYPHVLFDAQOTEPVIFTIPDIRKTLXH-SMDDTDTTLRLV 319  
QY 180 AMLYTPL-RSNGSGDDVFTVSCRVLTRTPDPFEITYLVPP-----SVESKTKP----- 226  
Db 320 IMVYNELINPVEQSPKSCSITVETRSDSFTSLAPPGSLLKHGSIPLSDILPRNSRH 379  
QY 227 -----FTLPILTIS-----ELTNSRPPIPIEOLYTAPNETNVVQONGRCT 267  
Db 380 WGNRNMSTIDGFFVQPRVFSQNRHDFDFTTGWSTP---YIPIEIVLEKLDRG--- 432  
QY 268 LDGELQGTQTLSSAVCFLOQRTVADNGNDWDLQLQTYNGASYDPTDEVAPLGTQD 327  
Db 433 --GQYFKVTDTEKSLVPLGLP-----DGMPTDTI-----PT-AMTASNGNYD 470  
QY 328 FSGMLYGLVLTQDNNVNVSTGEAKNAGIYISITSGKFTPKI-GSIGLHSITEHVHPNQSR 386  
Db 471 YVAEYRI-----TNGTHFGFYI---MGNLTTKVKGSDNLGETQQ-----TSRTL 514  
QY 387 FTPVGVADVENT 398  
Db 515 FASVGNVXKDQNT 526

RESULT 14

JQ2356

capsid protein - feline calicivirus (strain KCD)  
C:Species: feline calicivirus  
C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
C:Accession: JQ2356  
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.  
J:Gen. Virol. 74, 2519-2524, 1993  
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable regions  
A:Reference number: JQ2354; MUID:94065683; PMID:7504075  
A:Accession: JQ2356  
A:Molecule type: mRNA  
A:Residues: 1-568 <SEA>  
A:Cross-references: GB:L09719; NID:g305107; PIDN:AAAL6487.1; PID:g305108  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein

Query Match 9.4%; Score 268.5; DB 2; Length 668;  
Best Local Similarity 24.3%; Pred. No. 5.6e-12;  
Matches 132; Conservative 81; Mismatches 222; Indels 109; Gaps 27;  
QY 30 EPVAGASIAAPVAGQTNIIIDPW-IRTNFVQAPNGEFTVSPRNSPGBIILNLELGPDLNLY 88  
Db 145 EPNAQMSTAADMATCKSVDSWEAFSPHTSVNW---STSETQKILFKQSLGLNLY 200  
QY 89 LAHLSRMVYAGGVQVLLAGNAFTAGKILFAAIPPNFLVDMISPAQITMLPHLIVDV 148  
Db 201 LEHLAKLYVAVWSGSIEVRSISGSGVFGCLAAIIVVPEG--VDPVQSTSMQLQYPHVLFDA 258  
QY 149 RTLPIPTPLPDVNRVYHFNQPPQRMRLVAMLYTLP-----RSNGSGDDVFTVSCR 202  
Db 259 RQVFPVPSIPDLASTLYHLMSDITD-TTSLVMAYNDLINPYANDNSNGCIY-----TV 312  
QY 203 LTRTPDPFEITYLVPP-----SVESKTKPFTLPI-----LTISELTNSRFPPIEQ 248  
Db 313 ETKPGSDPRFHLKPPGSLVTHGSPVPSDLIPKTSLSLWGNRVYTNITNPFVIRPPVFQANR 372  
QY 249 LYTAPNETNVVQONGR-CTLDGELQGTQL-LSAVCFLOQRTVADNGDNW-DQLLQL 305  
Db 373 HFDFNQETAGWTPRFRPITITISEQGGTGLIGVATDYI----VPGIDGMPDPTTISEE 428  
QY 306 TYPNG-----ASYDPTDEVAPLGTQDFSGM-LYGLVLTQDNNVNVSTGEAK 349  
Db 429 LIPAGDVATINDIGNDITTPAGYDAADTIK---NNTNFRGMVYCGSLOR-----AWGDKK 480  
QY 350 NAKGIYISTT---SGKFTP-----KIGSIGLHSITEHVHPNQSR-----RFTPVG---V 392  
Db 481 ISNTAFITATVKNKLPKSNMIDQTKIAVFDQDNHVGDKVQTSDDTLALLGYTGIGEQAI 540  
QY 393 AVDENTPQQWVLPYHAGSLANTNLAPAVAPTPGEOQLLFRSRVPCVQGLQODAFID 452  
Db 541 GSDRDKVVRIISVLPETGAR-----CGNHPIFYKNSIKLVYVIRSIDVFN 585  
QY 453 CLL---PQWVNHFYQEAAPSQADVVALIRVYNPDGTLFEAKLHRSQ--FITVSHTG-- 505  
Db 586 QILTSRQLSLNHYL--LAPD--SFAVYRII--DSNGSWFDIGDSNGFSFGVSDIGKL 639  
QY 506 AYPL 509  
Db 640 EFPL 643

RESULT 15

C48562  
Query Match 9.4%; Score 268.5; DB 2; Length 668;  
Best Local Similarity 24.3%; Pred. No. 5.6e-12;  
Matches 132; Conservative 81; Mismatches 222; Indels 109; Gaps 27;  
N:Alternate names: capsid protein  
C:Species: San Miguel sea lion virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
C:Accession: C48562  
R:Neill, J.D.  
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lions.  
A:Reference number: A48562; MUID:92410750; PMID:1529644  
A:Accession: C48562

Wed Jun 2 09:13:35 2004

A:Molecule type: genomic RNA  
A:Residues: 1-703 <NEI>  
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888  
A>Note: sequence extracted from NCBI backbone (NCBIF:113567)  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 8.4%; Score 240; DB 1; Length 703;  
Best Local Similarity 29.3%; Pred. No. 88-10; Indels 16; Gaps 8;  
Matches 63; Conservative 45; Mismatches 91; Mismatches 16; Gaps 8;

QY 8 AAPSNDGAASLVPEG----INETMPLEPVAGASIAAPVAGQTNIIIDPWIRTNFVOAPNGE 63  
DB 151 AESDGPGEAEIVTEEGTVQQQPAPAPTALATLATASTGK-SVEQEW-TFF--SYHTS 206  
QY 64 FTVSPRNSPEILLNLELQPDNLNPNYLHLSRMVNGVAGGVEVQVLLAGNATAGKILFAA 123  
DB 207 INWSTVESQKILYQALNPSINPYLDHIAKLYSTWSGGIDVFTVSGSGVFGGKLAALL 266  
QY 124 IPPNFLVDMISPAQITMLPHLIVDVRTLSPIMTPLPDVRNVEYHFNNOPQPEMRLVAMLY 183  
DB 267 VPEG--VEPIESVSMQYPHVLFDAQRTPEVFTFTIPDIRKTLFHSNDETD-TTKLVINPY 323  
QY 184 TPLRSNGSGDDVFTVSCRVLTRPTDPDFEYLVPP 218  
DB 324 ----ENGV-ENKTTCSITVETREPSADFTFALLKPP 353

Search completed: June 1, 2004, 13:55:30  
Job time : 12.4581 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.6124 Seconds  
(without alignments)  
4985.230 Million cell updates/sec

Title: US-09-926-799-10  
Perfect score: 2906  
Sequence: 1 MKMASNDRAPSNDGAANLVP.....GNQFYTLAPMGSGQGRRAQ 550

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2020	69.5	542	S60616	capsid protein - h
2	1827.5	62.9	539	S40111	capsid protein - h
3	1141	39.3	530	B37471	capsid protein - N
4	1089.5	37.5	546	B37491	major capsid prote
5	319.5	11.0	576	A53982	capsid protein - E
6	293	10.1	702	A48562	coat protein - San
7	290	10.0	2344	RRWRH	genome polyprotein
8	288	9.9	2344	S55399	genome polyprotein
9	287.5	9.9	703	C48562	coat protein - San
10	284	9.8	2344	S64740	genome polyprotein
11	283	9.7	668	1 VCMWFC	coat protein - fel
12	280.5	9.7	671	1 VCMWFC	coat protein - fel
13	278.5	9.6	668	JQ2354	capsid protein - f
14	270.5	9.3	668	1 VCMWFF	capsid protein - f
15	262.5	9.0	668	JQ2356	capsid protein - f
16	129.5	4.5	2332	1 GNNYF	genome polyprotein
17	125.5	4.3	2194	1 GNNYE7	genome polyprotein
18	124.5	4.3	2333	1 GNNY2F	genome polyprotein
19	119.5	4.1	3085	2 T00327	polyprotein - infe
20	119	4.1	2647	2 A37098	gelation factor AB
21	118	4.1	2336	2 S37077	genome polyprotein
22	117.5	4.0	3375	2 T19821	hypothetical prote
23	116	4.0	733	2 JQ1892	capsid protein - f
24	115.5	4.0	490	2 T02411	cellulase (EC 3.2
25	115	4.0	733	2 JQ1891	capsid protein - f
26	111.5	3.8	2175	1 GNNYBE	genome polyprotein
27	111.5	3.8	2206	1 GNNY27	genome polyprotein
28	110.5	3.8	399	2 A11929	N-acetyl-glucosami
29	110.5	3.8	1204	2 S70393	pol polyprotein -

30	110.5	3.8	2332	1 GNNY4F	genome polyprotein
31	109.5	3.8	833	2 S45041	genome polyprotein
32	109.5	3.8	1199	1 GNMV1M	HIV-1 retropepsin
33	108.5	3.7	1204	2 S35475	pol polyprotein -
34	107	3.7	1011	1 GNNYC1	genome polyprotein
35	107	3.7	1658	2 D75489	hypothetical prote
36	106.5	3.7	614	2 S58306	WD-40 repeat regul
37	105	3.6	3566	1 A40701	tenascin-X procure
38	104.5	3.6	470	2 D84863	hypothetical prote
39	104.5	3.6	805	2 JC7635	aryl hydrocarbon r
40	104	3.6	492	2 T01584	cellulase (EC 3.2
41	104	3.6	3191	2 T22945	hypothetical prote
42	103.5	3.6	3624	2 AD0835	large repetitive p
43	103	3.5	1197	2 T30581	neural cell adhesi
44	102	3.5	883	2 T51221	hypothetical prote
45	102	3.5	1184	2 H70761	probable dnapolyme

ALIGNMENTS

RESULT 1

S60616 capsid protein - human calicivirus (strain Melksham)

C:Species: human calicivirus

A:Variety: strain Melksham

C:Date: 23-May-1997 #sequence\_revision 23-May-1997 #text\_change 28-Jul-2000

C:Accession: S60616

R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.

Virus Res. 37, 271-283, 1995

A:Title: Capsid diversity in small round-structured viruses: molecular characterization

A:Reference number: S60615; MUID:96136658; PMID:8533462

A:Accession: S60616

A:Molecule type: genomic RNA

A:Residues: 1-542 <GR>

A:Cross-references: EMBL:X81879; NID:G976077; PIDN:CAAS7462.1; PID:G976079

A:Experimental source: strain Melksham

A:Note: it is uncertain whether Met-1 or Met-3 is the initiator

C:Superfamily: human calicivirus capsid protein

C:Keywords: capsid protein; coat protein

Query Match 69.5%; Score 2020; DB 2; Length 542;

Best Local Similarity 69.3%; Pred. No. 5.1e-147;

Matches 383; Conservative 68; Mismatches 88; Indels 14; Gaps 5;

QY 1 MKMASNDRAPSNDGAANLVPPEANDEVMALPVGASIAAPVVGQNIIDPWIRENFVQAP 60

DB 1 MKMASNDRAPSNDGAAGLVPSNEVMALPVGAGALAAPVTGQNIIDPWIRANFVQAP 60

QY 61 QGEFTVSPRNSFGEMLLNLELGPENLPYLSHLSRMNGYAGGMQVQVVLGNAFTAGKII 120

DB 61 NGEFTVSPRNAPGVEVLLNLELGPENLPYLHLARMYNGYAGGMVQVVLGNAFTAGKLV 120

QY 121 FAAPVPHFPVENISAAQITMCPHVIVVQLEPVLPLPDINRFRFHNQENTPRLVA 180

DB 121 FAAPVPHFPVENISLPKTIETFPVHIIDVIRLEPVLPLPDVNRNSPFHYNKDDPMRIVA 180

QY 181 MLYTPLRAN-SEGEDVFTVSCRVLTPADPFRFTFLVPPVTESKTKPFTLPILTGLSLS 239

DB 181 MLYTPLRNSGDDVFTVSCRVLTPRPSDFFTLYLPPVTESKTKPFTLPILTGLSLS 240

QY 240 RFPAAIDMLYTDPNESIVVQNGRCITLDGLTQGTQTLVPTQICAFRGTLISQTAARADS 299

DB 241 RFPVPIDQMTSPNEVISVQNGRCITLDGLTQGTQTLQVSGICAFKGEV---TAHLHDN 297

QY 300 TDSQPARARNHLHVQVKMLDGTQDPTDDIPAVLGAIIDFKGTVFQVGSQRDVSQGQEQGH 359

DB 298 -----DHLNNVTITNLGSPEDPSDIPAPLGVDPFQGRVFGVI SQORDQKNAAGHSE 349

QY 360 YATRAHEAHIDTDPKYAPKLGTLIKS-GSDDFNTNPIRFTPVGMGD-NNRWQEWELPD 417

DB 350 PANRGHDAVPTTYTAQYTPKLGQIQIGTWQTDLLTVNQPVKFTPVGLNDTEHFNQWVPR 409

QY 418 YSGLTLLNNLAPAVSPSPGGRILFFRSIVPSAGGSGYIDCLIPQEWVQHFYQEAAP 477  
Db 410 YAGALNLTNLTAPSVAVPFPGERLLFFRSHLPKGGYGNPAIDCLLPQEWVQHFYQEAAP 469  
QY 478 SOSAVALVRYNPDTRGRNIFEAKLHREGFLTVANGCNPNPIVPPNGYFRFEANGNQFTTL 537  
Db 470 SNEVALVRYINPDTRGRALFEAKLHRAGMTVSSNTSAPVVVPFANGYFRFDSWVWQFYSL 529  
QY 538 APWGSQGGRRRAQ 550  
Db 530 APMTGNGRRRQVQ 542  
RESULT 2  
S40111  
capsid protein - human calicivirus (strain Bristol isolate B493)  
C;Species: human calicivirus  
A;Variety: strain Bristol isolate B493  
C;Date: 25-Dec-1994 #sequence\_revision 27-Feb-1997 #text\_change 28-Jul-2000  
C;Accession: S40111  
R;Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
submitted to the EMBL Data Library, December 1993  
A;Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de  
A;Reference number: S40111  
A;Accession: S40111  
A;Molecule type: genomic RNA  
A;Residues: 1-539 <GRE>  
A;Cross-references: EMBL:X76716; NID:G436410; PIDN:CAA54134.1; PID:G436411  
A;Experimental source: human enteric calicivirus strain Bristol isolate B493  
C;Superfamily: human calicivirus capsid protein  
C;Keywords: capsid protein; coat protein  
Query Match 62.9%; Score 1827.5; DB 2; Length 539;  
Best Local Similarity 62.1%; Pred. No. 3e-132;  
Matches 346; Conservative 83; Mismatches 101; Indels 27; Gaps 6;  
QY 1 MKMASNDAAPSNDGAA--NLVPEAN--DEVMALEPVVVGASIAAPVVGQOQNIIDPWIRENFVQAP 60  
Db 1 MKMASNDANPSDGAANLVPEANDEVMALEPVVVGAAIAAPVAGQOQNIIDPWIRNFVQAP 60  
QY 61 QGEFTVSPRSPGEMLLNLELGPENLPLSHLSRMVNGYAGGMQVQVVLGNAFTAGKII 120  
Db 61 GGEFTVSPRAPGELISAPLGPDLNPLSHLSRMVNGYAGGFEVQVILGNAFTAGKVI 120  
QY 121 FAAYPPHPFVENISAAQITMCPHVIVDVROLEPVLLPLDIRNPFHYNQENTPRMLVA 180  
Db 121 FAAYPPNPPTGLSPSQVTMPFHIIVDVROLEPVLLPLDIRNPFHYNQANDSTLKLJA 180  
QY 181 MLYTPLRA-NSGEDVFTVSCRVLTRPAPDPFETFLVPPTVESKTKPFTLPLITLIGLSNS 239  
Db 181 MLYTPLRANAGDDVFTVSCRVLTRPSPDPDFI FLVPTVESRTKPTVPLVTEENSNS 240  
QY 240 RFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLQGTTLQVLPQICAFRGTLISQTAADS 299  
Db 241 RFPIPLEKLYGPSSAFVVPQNGRCITTDGVLGTLTQLSAVNICNFRG----- 288  
QY 300 TDSPOARHPLHVQVKNLDGTQYDDTDIPAVLGAIKDFKTVGVSQORDVSQOQGH 359  
Db 289 -DVTHIAGSHDYTNLASQNSNYDPTBEIPAPLGTDPDFVKIQGLLT-----QTRAD 341  
QY 360 VATRAHEAHIDTDPKAPKLTILIKSG-SDDFNTNQPIRFTPVGM---GD---NNWRQ 412  
Db 342 GSTRAKATVSGSVHFTFKLGSVQFTDTNNDFOAGQNTKFTTPGVIGDQDHHQNEPQ 401  
QY 413 WELPDYSGRLTLNMLAPAVSPSPGGRILFFRSIVPSAGGSGYIDCLIPQEWVQHFY 472  
Db 402 WLLPNVSGRTHNVHLAPAVTPFGEQLLFFRSTMPGCSGYPNNNLDCLLPQEWVHLFY 461  
QY 473 QEAAPSQSAVALVRYNPDTRGRNIFEAKLHREGFLTVANGCNPNPIVPPNGYFRFEANGN 532  
Db 462 QBAAPQASQSAVALVRYNPDTRGRVLECKLHKSGLYITVAHTGTYDVLVLPNGYFRFDSMVN 521  
QY 533 QFYTLAPMGSGGRRRA 549

Db 522 QFYTLAPMGNGTGRRA 538  
RESULT 3  
B37471  
capsid protein - Norwalk virus  
C;Species: Norwalk virus  
C;Date: 24-Feb-1994 #sequence\_revision 25-Apr-1997 #text\_change 28-Jul-2000  
C;Accession: B37471  
R;Jiang, X.; Wang, K.; Wang, K.; Estes, M.K.  
Virology 195, 51-61, 1993  
A;Title: Sequence and genomic organization of Norwalk virus.  
A;Reference number: A37471; MUID:93303939; PMID:8391187  
A;Accession: B37471  
A;Status: preliminary; not compared with conceptual translation  
A;Molecule type: genomic RNA  
A;Residues: 1-530 <JIA>  
A;Cross-references: GB:M87661; NID:G1061311; PIDN:AAB50466.1; PID:G1061313  
A;Note: sequence extracted from NCBI backbone (NCBIP:134157)  
C;Superfamily: human calicivirus capsid protein  
Query Match 39.3%; Score 1141; DB 2; Length 530;  
Best Local Similarity 43.4%; Pred. No. 1.3e-79;  
Matches 249; Conservative 78; Mismatches 169; Indels 78; Gaps 12;  
QY 1 MKMASNDAAPSNDGAA--NLVPEAN--DEVMALEPVVVGASIAAPVVGQOQNIIDPWIRENF 56  
Db 1 MMMAKSDATSSVDGASGQQLVPEVNASDPLAMDVPVAGSSTAVATAGVNPIDPWIINNF 60  
QY 57 VQAPQGEFTVSPRSPGEMLLNLELGPENLPLSHLSRMVNGYAGGMQVQVVLGNAFTA 116  
Db 61 VQAPQGEFTISPNNTPGDVLFDLSLGPENLPLSHLSQMYNGWVGNMRVRLMLAGNAFTA 120  
QY 117 KLIIPAAVPPHPFVENISAAQITMCPHVIVDVROLEPVLLPLDIRNPFHYNQENTPRM 176  
Db 121 GKIIIVSCIPPGGSHNLTIQAATLFPFVIADVRLDPIEVLPLEVDRVNLFFNNDRNQTM 180  
QY 177 RLIVAMLYTPLRANSGE-DVFTVSCRVLTRPAPDPFETFLVPPTVESKTKPFTLPLITLGE 235  
Db 181 RLVCMLYTPRLTGGTGDSFVVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPLNPLSS 240  
QY 236 LNSRFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLQGTTLQVLPQICAFRGTLISQTA 295  
Db 241 LNSRAPLPISSIGISPDNVQSVQFQNGRCTLDGKLVGTTTPVSLSHVAKIRGT----- 293  
QY 296 AADSTDSPOARHPLHVQVKNLDGTQYDDTDIPAVLGAIKDFK-----TVFGVASQ 348  
Db 294 -----SNGTVINLTDELGTTPFHPPEG-PAPIGFDLGGCDWHNNMTQFGHSQ 340  
QY 349 RDVSGQOQGHVATRAHEAHIDTDPKAPKLTILIKS-QSDDFNTNQPIRFTPVGMGD 407  
Db 341 T-----QYDDVDTTPTTFVPHLGSIQANGISGNT-----VG 372  
QY 408 NNWRQ-----WELPDYSGRLTLNMLAPAVSPSPGGRILFFRSIVPSAGGY 455  
Db 373 LSWISPPSPSSQVDLWKIPNYGSIIEATHLASVYPGFGVLFVFMKMPGPGAYN 432  
QY 456 SGYIDCLIPQEWVQHFYQEAAPSQSAVALVRYNPDTRGRNIFEAKLHREGFLTVANGC-- 513  
Db 433 ---LPCLLQEQYISHLASEQAPTVGEAALLHYVDPDTRGNLGEFKAYPDGFLTCVPNGAS 489  
QY 514 NNPIVPPNGYFRFEANGNQFTLAPMGSGGRRR 547  
Db 490 SGPQQLPINGVFFVSVWSRFRYQLKPVGTASSAR 523  
RESULT 4  
B37491  
major capsid protein [similarity] - Southampton virus  
N;Alternate names: orf2 protein  
C;Species: Southampton virus  
C;Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 18-Aug-2000



C:Accession: B37491  
R:Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
Science 259, 516-519, 1993  
A:Title: Sequence and genome organization of a human small round-structured (Norwalk-like  
A:Reference number: A37491; MUID:93142023; PMID:8380940  
A:Accession: B37491  
A:Status: not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-546 <LAm>  
A:Cross-references: GB:L07418; NID:g1236787; PIDN:AAA92984.1; PID:g295114  
A>Note: sequence extracted from NCBI backbone (NCBIP:123458)  
A:Note: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype  
C:Superfamily: human calicivirus capsid protein  
C:Keywords: glycoprotein  
F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 37.5%; Score 1089.5; DB 2; Length 546;  
Best Local Similarity 42.6%; Pred. No. 1.2e-75;  
Matches 250; Conservative 75; Mismatches 173; Indels 89; Gaps 16;  
QY 1 MKAASNDAPSNDGAA---NLVPEAND-EVMALEPVVGASIAAPVVGQNIIDPWIRENF 56  
DB 1 MMASKADPQSDAGSAGQALPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPMVNF 60  
QY 57 VQAPQGEFTVSPRNGEMLNLELGPENLYSLSRMYNGYAGQVQVVLGNAFTA 116  
DB 61 VQSPQGEFTISPNTPGDILFDLQGLHLPFLSHLSQMYNGWGNMVRILLAGNAFSA 120  
QY 117 GKIIIPAAVPPHPPVENISAQITMCPHIVDVQLEPVLPLDPIRNRFFHYNQENTPRM 176  
DB 121 GKIIIVCCVPPGTSLSLTAAQLTFPHVJADVRLTEIEMPLEDVRNVDIHTN-DNQPTM 179  
QY 177 RLVMALYTLPRANS---EDVFTVSCRVLTRPAPDFEFTLVPPVTSKTKPFTLPILT 233  
DB 180 RLVMALYTLPRRTGGSGNSDSFVAGRVLTAPSSDFSLFVLPVPIEQKTRAFVPIPL 239  
QY 234 GELNSRPPAALDMLYTDPNESIVVQPNQGRCTLDGTLOGTTQVPTQICAPRGLISOT 293  
DB 240 QTLNSRFPSLIQMILSPDASQVQFQNGRCLIDQLLGLTTPATSGQLFRVRGK-INQG 298  
QY 294 ARAADSTSPQARNHPLHVQKNLDGTQYDTPDIPAVLGADF-----KG 340  
DB 299 AKTLNLT-----VDGKFFMAFDS-PAPVGPDPFGKCDWHRISKTPNN 341  
QY 341 TVFG-----VASQDVSGQEQGHYATRAHEAHIIDTPDKYAPKLGTLILKSGSDFN-- 393  
DB 342 TGGSDPMRSVSVQTNVQ-----FVPHLGSIQP---DEVFNHP 376  
QY 394 -----TNQPIRFTFPVGMGNRWQELPDYSGRLTLNMLAPAVSPSPGGERILF 443  
DB 377 TGDYIGTIEWISQP-STPPGTDIN---LWEIPDYGSSLSQAANLAPPVFPFGGEALVY 431  
QY 444 PRSIVPSAGGYS-GVIDCLIPQEWQHFYQQAAPSQSAVALRVYVNPDTGNIFEAKLH 502  
DB 432 FVSAPFGNNRRAPNDVPCLLPQEIYTHFVSQAPTMGDALHLYVDPDPTNRLNGFEKLY 491  
QY 503 REGFLTVA--NCCNNPIVVPNGYFRFEAWGNQFYTLAPMGSCQGR 547  
DB 492 PGGLTCTVNGVAGPQQLPLNGVFLFVSVWSRFFYQLKRPVGTASTAR 538  
RESULT 5  
A:Accession: A53982  
capsid protein - European brown hare syndrome virus  
C:Species: European brown hare syndrome virus  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 28-Jul-2000  
C:Accession: A53982  
R:Wirthlich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Eakens, U.; Haas, B.; Thiel, H.  
J. Virol. 68, 5164-5173, 1994  
A:Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease  
A:Reference number: A53982; MUID:94309183; PMID:7518531  
A:Accession: A53982  
A:Status: preliminary

A:Molecule type: mRNA  
A:Residues: 1-576 <MIR>  
A:Cross-references: GB:U09199  
C:Superfamily: human calicivirus capsid protein  
Query Match 11.0%; Score 319.5; DB 2; Length 576;  
Best Local Similarity 24.4%; Pred. No. 1.6e-16;  
Matches 136; Conservative 77; Mismatches 200; Indels 145; Gaps 22;  
QY 11 SNDGAANLVPEANDVEMALEPVVGASIAAPVVG-----QONIIDPWIRENFVQAPQGEFTV 66  
DB 22 ITDGDMDPGVASTDVVTA--DNVAASVATAGIGGPPQASQESW-RVNFFY--NDVFTW 76  
QY 67 SPRNSPGEMLNLELGPENLYSLSRMYNGYAGQVQVVLGNAFTAAGKIIFAAVPP 126  
DB 77 SVTDAPGSLYVSHSPQNNPFTQVLSQMYAGWAGQMPRFIVAGSGIFGRGLVCAIIPP 136  
QY 127 HPPVENISAQITMCPHIVDVQLEPVLPLDPIRNRFFHYNQENTPRMRLVAMLYTTL 186  
DB 137 GQIQI--FGLEVRQPFHVVIDARSLEPVTITMPDLRPEMYHPTGDPGLVTLVSYNNL 194  
QY 187 RANSGEDVFTVSCRVLTRPAPDFEFTLVPP---TVESKTKPF---TLPILTGLGELNSRF 241  
DB 195 INPFGTTSATQVTVETPSEDFEFLIRAPSSKTVDSVNPSSLTTPVLT--GAGSDNRW 253  
QY 242 PAADMLYTDPNESIVVQPNQGRCTLDGTLOGTTQ-----LVPTQICAPRGT 288  
DB 254 GAPIVGLQPVPGG---FSTSNRHNWNGSTYGSWSPRPDDIDHPSGNVSYPTG--SATNT 308  
QY 289 LISQTARAADSTDSQARARNHPLHVQKNLDGTQYDTPDIPAVLGADFPGKT----- 341  
DB 309 IETWYANAGTATNP-----ISNI-----APDGFDP-MGAIPFSGTTIPTGAW 350  
QY 342 -----VFGVASORDVS 352  
DB 351 VFGQVQVWASNGTPVGVTVQAYELGFANGAPSSIRPVTITTTGAQLVAKSIYGVATAQNS 410  
QY 353 GQ-----QEQGHYATRAHEAHIIDTPDKYAPKLGTLILKSGSD---DFNTNQPIRFTPV-- 403  
DB 411 SAGIIFLKSQVMVSTPGVAA---TT---YTPQPSAIVTTPGTVAAPIGKNTIMESAVVR 464  
QY 404 -----GMGDNWQWQELPDYSGRLTLNMLAPAVSPSPGGERILF-----FRSIVP 449  
DB 465 RTGDVNAAGSVNGTYGVGSGPLSLTLGSLTNYSALQPGQFFVQWLNFAFGFMVGM 524  
QY 450 SAGGY-----GSGYID 460  
DB 525 NTDGYFYAGTGAYSGMID 542

RESULT 6  
A:Accession: A48562  
coat protein - San Miguel sea lion virus (serotype 1)  
N:Alternate names: capsid protein  
C:Species: San Miguel sea lion virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
C:Accession: A48562  
R:Neill, J.D.  
Virus Res. 24, 211-222, 1992  
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel  
eins.  
A:Reference number: A48562; MUID:92410750; PMID:1529644  
A:Accession: A48562  
A:Molecule type: genomic RNA  
A:Residues: 1-702 <NEI>  
A:Cross-references: GB:M87481; NID:g334882; PIDN:AAA16217.1; PID:g334884  
A>Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBIF:113565)  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 10.1%; Score 293; DB 1; Length 702;  
Best Local Similarity 25.0%; Pred. No. 2.2e-14;

Matches 117; Conservative 61; Mismatches 150; Indels 140; Gaps 21;

QY 8 AAPSNDGAANLVPEANDEVMALPEVVGAS-----IAAPVVGQNIIDPWIRENFVQAPQGE 63

DB 151 AESDGGGADIVTEBGGTVQOQPVPAQSALTTLAAATG-----KTVDC 196

QY 64 FTV-----SPRNSGEMLLNLELGPPELNPYLHLSRMVNGYAGGMQVQVVLGNA 113

DB 197 WTTFSYHTAVNWSTEAGKILFSRALSPELNPYLHLSRSLYSSTWGGIDVRFVTSVSGS 256

QY 114 FTAGKIIFAAVPPHPPFVENISAAQITMCPHIVDVROLEPVLPLPDIRNRPFFHYNQENT 173

DB 257 VEGGKLAALIVPP--GIEVESPTMLQXPHVLFDARQTEPVFTIPDRIKTLIH--SMDDT 313

QY 174 PMRLVAMLYTPI--RANGSGEDVFTVSCRVLTRPAPDFEFTLVPTVESKTKPFTLPIL 231

DB 314 DTTIRLVIMYNELINIDYEOSEPKSCSIIVTRPSDDFTFLLKPP--GSLLKHSIPSD 371

QY 232 TLGELS-----NSRFPAAIDMLYDTPNESIVVQPO-----NGRCTLDGTGCTQLVPTQIC 283

DB 372 LIPRNSRHWMGNRWSTID-----GFVVQPRVFSNRHFDFTTTGWS--TFYXI- 420

QY 284 AFRGTLISQTARAADSTDFQPARNHLPHVQVKNLD-GTQYD-----P 325

DB 421 -----PIEVTLEKDRGGQYFKYVTTDEKSLVPLGPDGWP 454

QY 326 TDDIPAVL-----GAIDFKGTGVFVASQORDVSGQEQGHYATRAHEAHIIDTPDKYAPKLG 381

DB 455 DTTIPTAMTASNGYDYVTAEXRITN-----NGTHFKGFYI-----MG 492

QY 382 TLIIK-SGSDP-NTNQPIR--FTFVGMGNWNRQWELPDYSGLTLN 425

DB 493 NLATTVKVGSNLTGQTSTRTLSAVG-----NYKDQNTIN 528

RESULT 7

RRWRH

genome polyprotein - rabbit hemorrhagic disease virus

N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)

C;Species: rabbit hemorrhagic disease virus

C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998

C;Accession: A41039

R;Meyers, G.; Wirblich, C.; Thiel, H.J.

Virolgy 184, 664-676, 1991

A;Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing

A;Reference number: A41039; MUID:91361557; PMID:1840711

A;Accession: A41039

A;Molecule type: genomic RNA

A;Residues: 1-2344 <MEY>

A;Cross-references: GB:M67473

C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase

C;Keywords: nucleotidyltransferase

Query Match 10.0%; Score 290; DB 1; Length 2344;

Best Local Similarity 30.3%; Pred. No. 2.1e-13;

Matches 89; Conservative 49; Mismatches 114; Indels 42; Gaps 12;

QY 8 AAPSNDGAANLVPEANDEVMALPEVVGAS-----IAAPVVGQNIIDPWIRENFVQAPQGE 63

DB 1772 AAPQGEAAGTATTASVPGTTDGMDPGVVATTSVTAENSASATAGIGGPPQVQDOE 1831

QY 50 PWIRENFVQAPQGEFTVSPRNSPGEMLLNLELGPPELNPYLHLSRMVNGYAGGMQVQVVL 109

DB 1832 TW-RTNFY--NDVFTWSVADAPGSILYTVQHSPPQNPFATVLSQMYAGWAGMQFRFIV 1888

QY 110 AGNAFTAGKIIFAAVPPHPPFVENISAAQITMCPHIVDVROLEPVLPLPDIRNRPFFHYN 169

DB 1889 AGSGVFGGLVRAVIPP--GIEIGRGLVQRPFFHVVIDARSLEPVTITMPDLRPNMYHPT 1946

QY 170 QENTPRMLVAMLYTPIRANGSGEDVFTVSCRVLTRPAPDFEFTLVPP--TVESKTKP- 225

DB 1947 GDPGLVPTLVLSVYNNLINPFGGSAIQVTVETRPSEDFEFVMIAPSSKTVDSISPAG 2006

QY 226 -FTLPILTIGELSNRFPAAIDMLYDTPNESIVVQPOQNG---RCTLDGTGCTT 275

DB 2007 LTTTPVLT-GVGNDRWNGQI-----VGLQPVPGGFCNCRHWNLNGST 2049

RESULT 8

S55399

genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)

C;Species: rabbit hemorrhagic disease virus

A;Variety: isolate BS89

C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 18-Jun-1999

C;Accession: S55399

R;Rosai, C.

submitted to the EMBL Data Library, May 1995

A;Reference number: S55399

A;Accession: S55399

A;Molecule type: genomic RNA

A;Residues: 1-2344 <ROS>

A;Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641

A;Experimental source: isolate BS89

C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase

C;Keywords: polyprotein

Query Match 9.9%; Score 288; DB 2; Length 2344;

Best Local Similarity 30.4%; Pred. No. 3.1e-13;

Matches 89; Conservative 48; Mismatches 114; Indels 42; Gaps 12;

QY 9 APSNDGAANLVPEANDEVMALPEVVGAS-----IAAPVVGQNIIDPWIRENFVQAPQGE 63

DB 1773 AAPQGEAAGTATTASVPGTTDGMDPGVVATTSVTAENSASATAGIGGPPQVQDOE 1832

QY 51 WIRENFVQAPQGEFTVSPRNSPGEMLLNLELGPPELNPYLHLSRMVNGYAGGMQVQVVL 110

DB 1833 W-RTNFY--NDVFTWSVADAPGSILYTVQHSPPQNPFATVLSQMYAGWAGMQFRFIV 1889

QY 111 GNAFTAGKIIFAAVPPHPPFVENISAAQITMCPHIVDVROLEPVLPLPDIRNRPFFHYNQ 170

DB 1890 GSGVFGGLVRAVIPP--GIEIGRGLVQRPFFHVVIDARSLEPVTITMPDLRPNMYHPTG 1947

QY 171 QENTPRMLVAMLYTPIRANGSGEDVFTVSCRVLTRPAPDFEFTLVPP--TVESKTKP- 225

DB 1948 DGLVPTLVLSVYNNLINPFGGSAIQVTVETRPSEDFEFVMIAPSSKTVDSISPAGL 2007

QY 226 FTLPILTIGELSNRFPAAIDMLYDTPNESIVVQPOQNG---RCTLDGTGCTT 275

DB 2008 LTTTPVLT-GVGNDRWNGQI-----VGLQPVPGGFCNCRHWNLNGST 2049

RESULT 9

C48562

coat protein - San Miguel sea lion virus (serotype 4)

N;Alternate names: capsid protein

C;Species: San Miguel sea lion virus

C;Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999

C;Accession: C48562

R;Neill, J.D.

Virus Res. 24, 211-222, 1992

A;Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lions.

A;Reference number: A48562; MUID:92410750; PMID:1529644

A;Accession: C48562

A;Molecule type: genomic RNA

A;Residues: 1-703 <NEI>

A;Cross-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888

A;Note: sequence extracted from NCBI backbone (NCBIP:113567)

C;Superfamily: feline calicivirus coat protein

C;Keywords: capsid protein; coat protein; glycoprotein

F;89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 287.5; DB 1; Length 703;

Best Local Similarity 24.5%; Pred. No. 5.9e-14;

Matches 104; Conservative 67; Mismatches 144; Indels 109; Gaps 18;

```
QY      8 AAPSNDGAANLVPEANDEVMALEPVVG-----ASIAAPVVGQQNIIDPWIRENFVQAPQGE 63
Db      151 AESDGFSGSAEIVTEEQGTVVQQQPAPAPATLATATASTG-KSVEQEMWTFYSYHTSINW 209
QY      64 FTVSPRNSPGEMLLNLELGPENLPYLHLSRMVYAGGMQVQVVLGNAFTAGKLIIPAA 123
Db      210 STV-----ESGKILYSQALNPSINPYLDHIAKLYSTWSSGIDVRFTVSGSGVFGGKLAALL 266
QY      124 VPPHF--PVENISAAQITMCPHVIVDVRQLEPVLPLPDIRNFFHYNQENTPRMLIVAML 182
Db      267 VPPGVEPIESVMLOQ---PHVLFDARQTEPVIFITPIDIRKTLFH-SMDETDTTKLVINP 322
QY      183 YPLRANSGEDVFTVSCRVLTRPAPDFEFTFLVPP----- 217
Db      323 Y-----ENGVENKTCISITVETRPSADFTFALLKPPGSLIKHGSIFPSDLIPRNSAHMMGNR 378
QY      218 -----TVESKTKPFTLP---ILTLGEL-SNSRFPAAIDMLY 249
Db      379 WNSTISGFSVQPRVQSNRHFDFDSTTTCWSTPYVPYVPIBKIQKVGSKNKKWHVID---- 435
QY      250 TD-----PNESIVVQPNQGRCTLDGTLQ-GTTQLVPTQICA-FRGTLLIS 291
Db      436 TDKALVPGIGDPDPTTIDET---KATNGNFSYGESYRAGSTTIKPNENSTHFKGTIYC 492
QY      292 QTARAADSTDSQPARANHPLHVQKNLDGTQYDPTDDIPAVLGAIDFKGTVEFGVASQRDV 351
Db      493 GTLSTVEIPENDEQ---QIKTEAEKKSQTMVVT-----ADFKDTI--VKPQHKI 537
QY      352 SGQQ 355
Db      538 SPQK 541

RESULT 10
S64740
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
N:Contains: VP60 protein
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate AST/89
C:Date: 12-Jul-1996 #sequence revision 26-Jul-1996 #text_change 20-Jun-2000
R:Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
submitted to the EMBL Data Library, May 1995
A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by di
A:Reference number: S64740
A:Accession: S64740
A:Molecule type: genomic RNA
A:Residues: 1-2344 <CAS>
A:Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552
A:Experimental source: isolate AST/89
R:Boga, J.; Casais, R.; Martin-Alonso, J.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.
submitted to the EMBL Data Library, July 1993
A:Description: Molecular cloning, sequence and expression of the capsid protein gene fro
A:Reference number: S46944
A:Accession: S46944
A:Molecule type: genomic RNA
A:Residues: 1650-2344 <BOG>
A:Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623
A:Experimental source: isolate AST/89
R:Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp
A:Reference number: S49018; MUID:93255896; PMID:8488721
A:Accession: S49018
A:Molecule type: genomic RNA
A:Residues: 1650-1796 <PAR>
A:Cross-references: EMBL:Z24757
A:Experimental source: isolate AST/89
A:Accession: S65012
A:Molecule type: protein
A:Residues: 1767-1779;1875-1877,'X',1879-1881;1936-1938,'X',1940-1941 <PAW>
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: blocked amino end; polyprotein
```

```
Query Match          9.8%; Score 284; DB 2; Length 2344;
Best Local Similarity 30.4%; Pred. No. 6.2e-13;
Matches 89; Conservative 47; Mismatches 115; Indels 42; Gaps 12;

QY      9 AFSNDGAANL---VPEANDE-----VMALEPVV-----GASIAAPVVG-----QQNIIDP 50
Db      1773 APOGAAAGTATTASVPGTTTDCMDPGVWATTSVVTAENSASSIATAGIAGPPQVDDQOET 1832
QY      51 WIRENFVQAPQGEFTVSPRNSPGEMLLNLELGPENLPYLHLSRMVYAGGMQVQVVLGNAFTAGKLIIPAA 110
Db      1833 W-RTAFYV--NDVFTWSVADAFSGSLTYVQHSFQNNPFTAVLSQMYAGAGGMOFRFIVA 1889
QY      111 GNAFTAGKIIFAAPVPHPPVENISAAQITMCPHVIVDVRQLEPVLPLPDIRNFFHYNQ 170
Db      1890 GIGVFGRLVAAVIPP--GIEIGPGLVQRQPHVVIDARSLSPVTTITMDPLRPNMYHPTG 1947
QY      171 ENTPRMLVAMLYTPLRANSGEDVFTVSCRVLTRPAPDFEFTFLVPP---TVESKTKP-- 225
Db      1948 DFGLVPTLVLVSYNNLINPFGGTSIAIQTVETREPSDFEFYMIAPSSKTVDSISAGL 2007
QY      226 FTLPTLTGELSNRFPAAIDMLYDTPNESIVVQPNQ---RCTLDGTLQGT 275
Db      2008 LTPVLIT-GVGNDRWNGOI-----VGLQVPVGGFSTCNRHWNLNGST 2049

RESULT 11
VCWVFC
coat protein - feline calicivirus (strain Japanese F4)
N:Alternate names: capsid protein
C:Species: feline calicivirus
C:Date: 30-Jun-1992 #sequence_revision 30-Jun-1992 #text_change 16-Jun-2000
C:Accession: B40481
R:Tohya, Y.; Taniguchi, Y.; Takahaeahi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamazaki, Y.
Virology 183, 810-814, 1991
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.
A:Reference number: A40481; MUID:91306470; PMID:1853578
A:Accession: B40481
A:Molecule type: genomic RNA
A:Residues: 1-668 <TOH>
A:Cross-references: GB:D90357; NID:g221264; PIDN:BAAL4371.1; PID:g221266
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:177,301,304,399,459,615/Binding site: carbohydrate (Aan) (covalent) #status predicted

Query Match          9.7%; Score 283; DB 1; Length 668;
Best Local Similarity 24.2%; Pred. No. 1.2e-13;
Matches 131; Conservative 78; Mismatches 198; Indels 134; Gaps 27;

QY      11 SNDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQQNIIDPWIRENFVQAPQGEFTVSP 68
Db      125 ADDGSIT-TPEQGTAVGVIAEPSAQMTAADMASGKSVDSW--EAFPSF-HTSVNWS 180
QY      69 RNSPGEMLLNLELGPENLPYLHLSRMVYAGGMQVQVVLGNAFTAGKLIIFAAPVPHF 128
Db      181 SETQKILFKQSLGPLNPLYLHLSKLYAVAWSGSEIVRFSISGSGVFGKLAALIVVPPGV 240
QY      129 -PVENISAAQITMCPHVIVDVRQLEPVLPLPDIRNFFHYNQENTPRMLVAMLYTPL- 186
Db      241 DPVQSTSMLOQ---PHVLFDARQVEPVIFITPIDLRSTLYHV-MSDITDTSLSVIMYNDLI 296
QY      187 -----RANSGEDVFTVSCRVLTRPAPDFEFTFLVPPPTVESKTKPFTLPTLTGELSNR 241
Db      297 NPYANDSNSGCIIVTE---TKPGDPDFKHLKPPG-----SVLTHGSPSOLI 342
QY      242 PAADIM-----LYTDPNESIVVQPNQGRCTLDGTLQGTQLVPTQICAFRGTLLISQTARA 296
Db      343 PKSSSLWIGNRYWTDITD-FVIRP-----FVFAQNRH 373
QY      297 AD---STDSPQARNHPLHVQKNLDGTQYDPTDDIPAVLGAIDFKGTVEFGVASQRDVSG 353
Db      374 FDFNQETAGWSTPRPRPRTITISEKNGSK-----LG-----IGVATDYIIPG 415
```

354 QOEQGHYATRAHEAHTDTPKAPKL---GTLIKSG-SDDFNTNQPIRFTPVGMGNN 409  
 416 -----IPDGHFDIT---IADKLIPAGDYSITTEGNDIKTAQAYDTAAVVKNTN 462  
 410 WRQWELPDYSGRLTNMNLAPAVSPFGRILFRSIVPSAGVG-----SGYID--C 461  
 463 PR-----GMVIGSLQAW-GDKKISNTAFITTAIRDNGEIKPSNTIDMTKL 508  
 462 LIPOEWQHFYOEAAPSQSAVALRVNPDTCRNIPEAKLHREGFLTV-----ANCNNP 516  
 509 AYQD--THVEQEVQTSDDTLALLGVTG--IGEEAIGSNRDEVRISVLPEAGARGGNHP 564  
 517 I 517  
 565 I 565

RESULT 12  
 VCWVF9  
 coat protein - feline calicivirus (strain F9)  
 N:Alternate names: capsid protein  
 C:Species: feline calicivirus  
 C:Date: 30-Jun-1993 #sequence revision 30-Jun-1993 #text\_change 28-Jul-2000  
 C:Accession: B43382; #sequence revision 30-Jun-1993 #text\_change 28-Jul-2000  
 R:Cartier, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.  
 Virol. 190, 443-448, 1992  
 A:Title: The complete nucleotide sequence of a feline calicivirus.  
 A:Reference number: A43382; MUID:92410623; PMID:1529544  
 A:Accession: B43382  
 A:Molecule type: genomic RNA  
 A:Residues: 1-671 <CAR1>  
 A:CROSS-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
 R:Cartier, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.  
 Arch. Virol. 122, 223-235, 1992  
 A:Title: Identification and sequence determination of the capsid protein gene of feline  
 A:Reference number: A45538; MUID:92117861; PMID:1731695  
 A:Accession: C45538  
 A:Molecule type: genomic RNA; protein  
 A:Residues: 1-671 <CAR2>  
 A:CROSS-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
 A:Experimental source: strain F9  
 A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBI:77462)  
 R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.  
 J. Gen. Virol. 73, 2429-2433, 1992  
 A:Title: The cloning, sequencing and expression of a major antigenic region from the fel  
 A:Reference number: PQ0407; MUID:93019069; PMID:1402818  
 A:Accession: PQ0407  
 A:Status: nucleic acid sequence not shown  
 A:Molecule type: mRNA  
 A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>  
 A:CROSS-references: PIDN:AA823553.1; PID:g257083  
 C:Superfamily: feline calicivirus coat protein  
 C:Keywords: capsid protein; coat protein; glycoprotein  
 F:177,304,439,459,618/Binding site: carbohydrate (Aen) (covalent) #status predicted

Query Match 9.7%; Score 280.5; DB 1; Length 671;  
 Best Local Similarity 24.1%; Pred. No. 1.9e-13;  
 Matches 130; Conservative 74; Mismatches 207; Indels 129; Gaps 24;

11 SNDGAANLVPEANDEV--MALEPVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68  
 125 ADDGSIT-APEQGTWVGVIAPSAQMSAAMATGKSVDSW--EAFPSF-HTSVNMST 180  
 69 RNSPGEMLNLBELGELNPNYLSHLSRMVNGVAGGQVQVVLGNAFTAGKIIFAAVPPHF 128  
 181 SETQKILFKQSLGPLNLPYLSHLAKLYVAMSGSVFVSISGSGVFGGKLAIVVPPGV 240  
 129 -PVENISAAQITMCHVIVDVQLEPVLPLPDINRFRFHYNOENTPRMRLVAMLYTPL- 186  
 241 DPVQSTSMQY---PHVLFDARQVEVFCLEPDLASTLYHL-MSDITSLVIMVNDLI 296  
 187 -----RANGSDVFTVSCRVLTPAPDFEFTFLVPVTSKTPFTLILTIGELSNRFF 241

297 NPYANDANSSGCIIVTVE-----TKPDPDFKHLKKPPG-----SMLTHGSTPSDLI 342  
 242 PAA-----IDMLYTDPNESIVVQPNQGRCTLDGTLQTTQLVPTQICAFRGTILISQTARAA 297  
 343 PKTSSLLWGNRYWSDITDFVIRP-----FVQANRHF 374  
 298 D---STDSPORARNHPLHVQKNLDCGTQYDPTDDIPAVLGAIDFKGTVFVASORDVSQ 354  
 375 DFNQETAGSTPFRPISVITTEQNG-----AKLG-----IGVATDYIVPG- 415  
 355 QEQGHYATRAHEAHTDTPKAPKLGTILIKSGS--DDFNTNQPIRFTPVGMGNNRWQ 413  
 416 -----IPDGPWPTTIPGELIPAGDYAITNGTNDITATGATYDADIKNNTFR-- 464  
 414 ELPDYSGRLLTNMNLAPAVSPSPGRILFRSIVPSA---GGYSGYIDC-LIPOEWV- 468  
 465 -----GMVIGSLQAW-GDKKISNTAFITATLDGNNKINPCNTIDQSKIV 512  
 469 ---QHFOEAAPSQSAVALRVY-----VNPDTGRNIFPEAKLHREGFLTVANCGNPI 517  
 513 VFQDNHVGGKKAQTSDDTLALLGYTGIGEQAIGSDRDRVVRISTLPETG-----ARGGNHPI 568

RESULT 13  
 JQ2354  
 capsid protein - feline calicivirus (strain NADC)  
 C:Species: feline calicivirus  
 C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
 C:Accession: JQ2354  
 R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.  
 J. Gen. Virol. 74, 2519-2524, 1993  
 A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable  
 A:Reference number: JQ2354; MUID:94065683; PMID:7504075  
 A:Accession: JQ2354  
 A:Molecule type: mRNA  
 A:Residues: 1-668 <SEA>  
 A:CROSS-references: GB:L09718; NID:g305104; PIDN:AAA16485.1; PID:g305105  
 C:Superfamily: feline calicivirus coat protein  
 C:Keywords: capsid protein; coat protein

Query Match 9.6%; Score 278.5; DB 2; Length 668;  
 Best Local Similarity 24.1%; Pred. No. 2.7e-13;  
 Matches 125; Conservative 73; Mismatches 193; Indels 127; Gaps 22;

30 EPVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSPRNSPGEMLNLELGPENLYL 89  
 145 EFSAQMSAAMATGKSVDSW--EAFPSF-HTSVNMSTSETQKILFKQSLGPLNLYL 201  
 90 SHLSRMVNGVAGGQVQVVLGNAFTAGKIIFAAVPPHF-PVENISAAQITMCHVIVDV 148  
 202 EHLKLYVAMSGSVFVSISGSGVFGGKLAIVVPPGVDPVQSTSMQY---PHVLFDA 258  
 149 ROLPEVLLPLPDINRFRFHYNOENTPRMRLVAMLYTPL-----RANGSDVFTVSCRVL 202  
 259 RQVDPVIFIPDLASTLYHL-MPDITSLVIMVNDLINPVANDSNSSGCIIVTVE---- 313  
 203 TRPADPFTFLVPVTSKTPFTLILTIGELSNRFPAA-----IDMLYTDPNESIVV 258  
 314 TKRGPDKFHLKKPPG-----SMLTHGSVPSDLIPKSSSLWGNRYWSDITDFV 363  
 259 QPNQGRCTLDGTLQTTQLVPTQICAFRGTILISQTARAA---STDSPORARNHPLHVQV 315  
 364 RP-----FVQANRHFNFQETAGMSAPRFRPITITI 395  
 316 KNLDGTQYDPTDDIPAVLGAIDFKGTVFVASORDVSQEQGHYATRAHEAHTDTPDK 375  
 396 SESKGSK-----LG-----IGVATDYIVPG-----IPDGPWPTTIAE 427  
 376 YAPKLTILIKSGS--DDFNTNQPIRFTPVGMGNNRWQELPDYSGRLTNMNLAPAVSP 434  
 428 DLTPAGDYAITSGNGNDITGSEYDSTEVIKNTNFR-----GMVIGSLQ 474  
 435 SFGPERILFRSIVPSAGVG-----SGYID---CLIPQEWQHFYOEAAPSQSAVALVR 486

Db 475 AW-GDKKISNTAFITAIKEGNKIRPSNTIDMTKIAYQD--THVGEVQTSDDALALIG 531  
QY 487 Y-----VNPTGRNIFEAKLHREGFLTVANGNNPI 517  
Db 532 YTGIGEQAIGSDRDRVVRISVLPEVG----ARGGNHPI 565

RESULT 14  
VCWVFF  
coat protein - feline calicivirus (strain CFI/68 FIV)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: A40507; B40507; T09246  
R:Neill, J.D.; Reardon, I.M.; Heinrichson, R.L.  
J. Virol. 65, 5440-5447, 1991  
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calicivirus  
A:Reference number: A40507; MUID:91374597; PMID:1716692  
A:Accession: A40507  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <NE1>  
A:Cross-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875  
A:Accession: B40507  
A:Molecule type: protein  
A:Residues: 373-379;403-419;481-489;560-566 <NE2>  
R:Neill, J.D.  
A:Submitted to the EMBL Data Library, April 1998  
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.  
A:Reference number: Z16626  
A:Accession: T09246  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <NE3>  
A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256  
A:Experimental source: strain CFI/68 FIV  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.3%; Score 270.5; DB 1; Length 668;  
Best Local Similarity 31.1%; Pred. No. 1.1e-12;  
Matches 75; Conservative 41; Mismatches 94; Indels 31; Gaps 10;

QY 11 SNDGAANLVPEANDEV--MALEPVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68  
Db 125 ADDGSI-TPEQGTWVGVIAPNAQMSTADMATGKSVDSW--EAFSP-HTSVNWS 180  
QY 69 RNSPGEMLLNLELGPENLPYLHLAKLYVWSGSVDRFSISGSGVFGGKLAIVVPPGI 240  
Db 181 SETQKILFKQSLGPLNLPYLHLAKLYVWSGSVDRFSISGSGVFGGKLAIVVPPGI 240  
QY 129 -PVENISAAQITMCPHVIVDVRLQLEPVLPLPDINRPFHYNQENTPRMLVAMLYTLP- 186  
Db 241 DPVQSTSMQY---PHVLFDARQVPEVIFSPDLASTLYHL-MSDITDTTSLVIMVYNDLI 296  
QY 187 -----RANGSEDVFTVSCRVLTRPADPEFTFLVPPTVESKTPFTLPILTGLGELNSRF 241  
Db 297 NPYANDSNSGCIIVTE----TKPGDFKHLKPPG-----SNLTHGSIPLSDLI 342  
QY 242 P 242  
Db 343 P 343

RESULT 15  
JQ2356  
capsid protein - feline calicivirus (strain KCD)  
C:Species: feline calicivirus  
C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
C:Accession: JQ2356  
R:Seal, B.S.; Kidpath, J.F.; Mengeling, W.L.  
J. Gen. Virol. 74, 2519-2524, 1993

A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable regions  
A:Reference number: JQ2354; MUID:94065683; PMID:7504075  
A:Accession: JQ2356  
A:Molecule type: mRNA  
A:Residues: 1-668 <SEA>  
A:Cross-references: GB:L09719; NID:g305107; PIDN:AAA16487.1; PID:g305108  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein

Query Match 9.0%; Score 262.5; DB 2; Length 668;  
Best Local Similarity 30.7%; Pred. No. 4.6e-12;  
Matches 74; Conservative 40; Mismatches 96; Indels 31; Gaps 10;

QY 11 SNDGAANLVPEANDEV--MALEPVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68  
Db 125 ADDGSI-TPEQGTWVGVIAPNAQMSTADMATGKSVDSW--EAFSP-HTSVNWS 180  
QY 69 RNSPGEMLLNLELGPENLPYLHLAKLYVWSGSVDRFSISGSGVFGGKLAIVVPPGV 240  
Db 181 SETQKILFKQSLGPLNLPYLHLAKLYVWSGSVDRFSISGSGVFGGKLAIVVPPGV 240  
QY 129 -PVENISAAQITMCPHVIVDVRLQLEPVLPLPDINRPFHYNQENTPRMLVAMLYTLP- 186  
Db 241 DPVQSTSMQY---PHVLFDARQVPEVIFSPDLASTLYHL-MSDITDTTSLVIMVYNDLI 296  
QY 187 -----RANGSEDVFTVSCRVLTRPADPEFTFLVPPTVESKTPFTLPILTGLGELNSRF 241  
Db 297 NPYANDSNSGCIIVTE----TKPGDFRHLKPPG-----SVLTHGSIPLSDLI 342  
QY 242 P 242  
Db 343 P 343

Search completed: June 1, 2004, 13:55:32  
Job time : 12.6124 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 46.8188 Seconds  
(without alignments)  
3289.030 Million cell updates/sec

Title: US-09-926-799-1  
Perfect score: 2896  
Sequence: 1 MMASKADPTNMDGTSGAGQ.....YOLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2896	100.0	545	4	AAB49700 Small rou
2	1979	68.3	544	4	AAB49703 Small rou
3	1977.5	68.3	546	4	AAB49702 Small rou
4	1950	67.3	530	2	AAR50972 Norwalk v
5	1950	67.3	530	7	ADC72176 Norwalk v
6	1938	66.9	530	4	AAB49701 Small rou
7	1900	65.6	530	2	AAR57091 Small rou
8	1225	42.3	540	4	AAB49706 Small rou
9	1178.5	40.7	548	4	AAB49705 Small rou
10	1174.5	40.6	535	4	AAB49707 Small rou
11	1172.5	40.5	542	4	AAB49708 Small rou
12	1157.5	40.0	539	4	AAB49704 Small rou
13	1157	40.0	548	5	AAU91272 Norwalk v
14	1154.5	39.9	550	4	AAB49709 Small rou
15	1013	35.0	541	4	AAB49710 Small rou
16	321	11.1	579	2	AAW08143 RHDV caps
17	276	9.5	547	4	AAW50108 Feline ca
18	276	9.5	671	4	AAW50107 Feline ca
19	272.5	9.4	669	4	AAB67461 Amino aci
20	272	9.4	623	4	AAB47044 Feline ca
21	272	9.4	668	4	AAB67462 Amino aci
22	269	9.3	623	4	AAB47043 Feline ca
23	269	9.3	668	2	AAR10886 Feline ca
24	269	9.3	668	4	AAE04304 Feline ca
25	262.5	9.1	622	4	AAB47045 Feline ca

26	154	5.3	2206	2	AAR22210	Aar22210 True type
27	147	5.1	40	5	AAU91273	Aau91273 Norwalk v
28	136	4.7	40	5	AAU91274	Aau91274 Norwalk v
29	132.5	4.6	2209	1	AAP20037	Aap20037 Sequence
30	125.5	4.3	2179	1	AAP60243	Aap60243 Sequence
31	122.5	4.2	6310	6	ABU39869	Abu39869 Protein e
32	122	4.2	3263	6	ABU49739	Abu49739 Protein e
33	116.5	4.0	1707	4	ABG22165	Abg22165 Novel hum
34	114.5	4.0	1765	6	ABU20131	Abu20131 Protein e
35	113	3.9	1555	6	ABJ25640	Abj25640 Aspergill
36	113	3.9	1832	6	ABJ26240	Abj26240 Aspergill
37	112.5	3.9	2164	1	ABP81045	Abp81045 Sequence
38	112.5	3.9	2164	1	ABP80131	Abp80131 Peptides
39	112.5	3.9	2914	5	ABP70089	Abp70089 Human NOV
40	112	3.9	423	7	ABR82681	AbR82681 Human mit
41	112	3.9	3930	6	ABU18893	Abu18893 Protein e
42	110.5	3.8	5291	7	ADC01014	Adc01014 Enterohae
43	110	3.8	1037	4	ABH71260	Abh71260 Drosophil
44	110	3.8	1234	4	ABH68510	Abh68510 Drosophil
45	110	3.8	2016	4	ABB63911	Abb63911 Drosophil

ALIGNMENTS

RESULT 1  
AAB49700  
ID AAB49700 standard; protein; 545 AA.  
XX  
AC AAB49700;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 1.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX  
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
XX  
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX  
WPI, 2001-080848/09.  
XX  
N-FSDB; AAF29141.  
XX  
Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.  
XX  
Claim 1; Page 40-42; 84pp; Japanese.  
XX  
This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks  
XX  
SQ Sequence 545 AA;

Query Match 100.0%; Score 2896; DB 4; Length 545;  
Best Local Similarity 100.0%; Pred. No. 9,7e-250;





OS Small round structured virus.  
 XX WO200079280-A1.  
 XX PD 28-DEC-2000.  
 XX PF 22-JUN-2000; 2000WO-JP004095.  
 XX PR 22-JUN-1999; 99JP-00175928.  
 XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 XX PA (DENK-) DENKA SEIKEN KK.  
 XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 DR WPI; 2001-080848/09.  
 DR N-PSDB; AAF29143.  
 XX Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.  
 XX Claim 1; Page 45-47; 84pp; Japanese.  
 CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX Sequence 546 AA;  
 SQ  
 Query Match 68.3%; Score 1977.5; DB 4; Length 546;  
 Best Local Similarity 66.6%; Pred. No. 1.3e-167;  
 Matches 367; Conservative 76; Mismatches 95; Indels 13; Gaps 7;  
 1 MMASKADPTNMDTSGAGQLVPEANTAEPTSMPEVAGAAATAAGOVNMDPWIMNMY 60  
 1 MMASKADQASDAGSAGQLVPEVNTADPLMEPVAGPTTAVATAGOVNMDPWVNNF 60  
 61 VQAPQGEFTISPNNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120  
 61 VQSPQGEFTISPNNTPGDILFDLQGLPHNLPFLSHLSQMYNGWGNMVRILLAGNAPSA 120  
 121 GKIIISCIIPCGFAAQNISIAQATMFPVIAADVRLVLEPLEVLEPVDVRLVFNHNDNAPTMR 180  
 121 GKIIIVCCVPPGFTSSSLTIAQATLFPVIAADVRLVLEPLEVLEPVDVRLVFNHNDNAPTMR 180  
 181 LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTPFSPVNLPLN 240  
 181 LVCMLYTPLRTGGSGNSDSFVAGRVLTAFSSDFSLFLVPPVPTIEQKTRAFVNPILQ 240  
 241 TLSNRSVPSLIKSMVSRDHQGMVQFNGRVTLQGLQGTTPTSASQCKIRGVSFVHANG 300  
 241 TLSNRSFSLIQGMILSPDASQVQFNGRCLIDQGLLGTTPATSGQLFRVRGKI--NQG 298  
 301 GNGVNLTLDCSPHAFESPAPIGPDPDLGECDDHMEASPT--TQNTGTVIQLINKVQE-S 358  
 299 ARTNLNLTEDGKPFMAFSPAPVGPDPFGKCDHMRISKTNNNTSSGDMRSVSQTNVQ 358  
 359 AFAPHLGTIQADGLSDVSNNTMLAKLWVSPVSDGHRGDDVDPVPIRYGSLTLEAAQLA 418  
 359 GFVPHLSIQDEVPNHGTG-DYGTIEWISQPTTPGCTDNLMEIPDYGSSLSQANLA 417  
 419 PPIVPPGGEAIVFMSDFPIAHGNGLS----VPTTIPQBFVTHFVNEQAPTGEAALL 474  
 418 PPVFPFGGEALVYFVSAPP--GNPNSAPNDVPCLLPQBYITHFVSEQAPTMGDAALL 474  
 475 HYLDPDTHNLGKFLYPEGFWTCVPNSSGTGPQTLPIGNGVYFVSVWSRYQLKPVGTA 534  
 475 HYVDPDTHNLGKFLYPEGFWTCVPNSSGTGPQTLPIGNGVYFVSVWSRYQLKPVGTA 534

QY 535 GPA-CRLGIRR 544  
 DB 535 STARSRLGVR 545  
 RESULT 4  
 AAR50972  
 ID AAR50972 standard; protein; 530 AA.  
 XX AC AAR50972;  
 XX DT 16-OCT-2003 (revised)  
 XX DT 25-MAR-2003 (revised)  
 XX DT 05-OCT-1994 (first entry)  
 XX DE Norwalk virus strain 8FIIa protein (encoded by ORF2).  
 XX Norwalk virus; pathogen; acute gastroenteritis; food poisoning;  
 KW seafood contamination; diagnostic assay; calcivirus; small round virus.  
 XX Norwalk virus; (strain 8FIIa).  
 XX WO9405700-A2.  
 XX PD 17-MAR-1994.  
 XX PF 07-SEP-1993; 93WO-US008447.  
 XX PR 07-SEP-1992; 92US-00941365.  
 XX PA (BAYU) BAYLOR COLLEGE MEDICINE.  
 XX PI ~~MASSON DOG, ESTES MK, Ulang, Graham, DY~~  
 XX WPI; 1994-101125/12.  
 DR N-PSDB; AAQ56826.  
 DNA from Norwalk and related viruses - used for preparing prods. for use  
 in diagnostic assays, detection and vaccines for Norwalk and related  
 viruses.  
 Claim 14; Page 68-70; 156pp; English.  
 The Norwalk virus was isolated from stool samples from adult volunteers  
 infected with safety tested Norwalk virus strain 8FIIa. The coding  
 sequence is useful for the design of probes for use in diagnostic assays  
 for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct  
 PN field.) (Updated on 16-OCT-2003 to standardise OS field)  
 SQ Sequence 530 AA;  
 Query Match 67.3%; Score 1950; DB 2; Length 530;  
 Best Local Similarity 67.0%; Pred. No. 3.5e-165;  
 Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;  
 1 MMASKADPTNMDTSGAGQLVPEANTAEPTSMPEVAGAAATAAGOVNMDPWIMNMY 60  
 1 MMASKADTSSVDASGAGQLVPEVNASDPLAMPVAGSTAVATAGOVNPDWVNNF 60  
 61 VQAPQGEFTISPNNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120  
 61 VQAPQGEFTISPNNTPGDILFDLQGLPHNLPFLSHLSQMYNGWGNMVRILLAGNAFTA 120  
 121 GKIIISCIIPCGFAAQNISIAQATMFPVIAADVRLVLEPLEVLEPVDVRLVFNHNDNAPTMR 179  
 121 GKIIIVCCVPPGFTSSSLTIAQATLFPVIAADVRLVLEPLEVLEPVDVRLVFNHNDNAPTMR 180  
 180 RLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTPFSPVNLPL 239  
 181 RLVCMLYTPLRTGGGTG--DSFVAVAGRVMTCPSPDFSLFLVPPVQKTPFSLPLNPL 238  
 240 NTLNSRVPSLIKSMVSRDHQGMVQFNGRVTLQGLQGTTPTSASQCKIRGVSFVHANG 299

```

Db      239  SSLSNRAPLPISSMGISPDNVQVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT-----S 294
Qy      300  GNGYNLTDLGSPYHAFESAPIGFDPDLGECMDHMEASPTQNTGDKVTKQINVKQESA 359
Db      295  NGTVINLTDLGTPFHPEGPAPIGFDPDLGECMDHIN---MTQFGHSSQTYQYDVTTPDT 351
Qy      360  FAPHLGTIQADGLSDVSNTNMIKLGWVSPVSDGHRGVDVPMWIPRYGSTLTTEAAQLAP 419
Db      352  FVPHLGSIQANGIG-----SGNYVGVLSWISPPSHPSGSQVDLWKIPNYGSSITEATHLP 407
Qy      420  PIYPGFGGALVFFMSDPPIAHNGTNGLSVPCITIOEFVTHFVNECAPTRGEAALLHYLDP 479
Db      408  SVYPPGFGELVFFMSKMP---GPGAYNLPCLLPQEIYISHLASQAPTGVGEAALLHYVDP 464
Qy      480  DTHRNLGFEKLYPEGFMTCVNSSGTGPQTLPIGVFVSVWSRFYQKLPVGTAGPA-C 538
Db      465  DTGRNLGFEKAYPDGFLTCVNGASSGQQLPIGVFVSVWSRFYQKLPVGTASSARG 524
Qy      539  RLGIIR 544
Db      525  RLGLRR 530

RESULT 5
ADC72176
ID      ADC72176 standard; protein; 530 AA.
XX
AC      ADC72176;
XX
DT      18-DEC-2003 (first entry)
XX
DE      Norwalk virus protein 2 amino acid sequence.
XX
KW      immune response; non-Norwalk virus agent; immunogen; Norwalk virus;
KW      viral pathogen; acute gastroenteritis; virucidal; antiviral vaccine.
XX
OS      Norwalk virus.
XX
PN      US6572862-B1.
XX
PD      03-JUN-2003.
XX
PF      07-JUN-1995; 95US-00486049.
XX
PR      08-NOV-1989; 89US-00433492.
XX
PR      27-APR-1990; 90US-00515993.
XX
PR      27-AUG-1990; 90US-00573509.
XX
PR      06-MAY-1991; 91US-00696454.
XX
PA      (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI      Gates MK, Jiang X, Graham DY;
XX
DR      MPI; 2003-776005/73.
XX
DR      N-PSDB; ADC72174.
XX
XX
XX      Inducing an immune response against non-Norwalk virus agents, comprises
XX      administering an immunogen recombinantly expressed from a cDNA from
XX      Norwalk virus.
XX
XX      Example 4; SEQ ID NO 3; 45pp; English.
XX
XX      This invention relates to a novel method of inducing an immune response
XX      in an individual against Norwalk virus and non-Norwalk virus agents, by
XX      orally or parenterally administering an immunogen recombinantly expressed
XX      or synthesised from a cDNA of Norwalk virus given in the specification.
XX      Norwalk virus is one of the most important viral pathogens, causing acute
XX      gastroenteritis. The invention may be used for the development of
XX      compounds with virucidal activity or an antiviral vaccine. The present
XX      sequence is the amino acid sequence of a protein encoded by the Norwalk
XX      virus genome of the invention.
XX

```

```

SQ      Sequence 530 AA;
Query Match      67.3%; Score 1950; DB 7; Length 530;
Best Local Similarity 67.0%; Pred. No. 3.5e-165;
Matches 366; Conservative 66; Mismatches 96; Indels 18; Gaps 7;

Qy      1  MMASKADAPTNMDGTSGAGQLVPEANTAEPIISMPEFVAGATAAATAGQVNMIDPMNMY 60
Db      1  MMASKADATSSVDGASGAGQLVPEYNASDPLAMDPAVAGSTAVATAGQVNPIDPMIINF 60
Qy      61  VQAPQGEFTISPNPTGDIILFDLQGLPHNLPLSLHAQMYNGWGVGMKVVLGAGNAFTA 120
Db      61  VQAPQGEFTISPNPTGDIILFDLQGLPHNLPLSLHAQMYNGWGVGMKVVLGAGNAFTA 120
Qy      121  GKIISCIIPGPAQNIATAQMTPEPHVTADVRVLEPIEVPLEDVRNVLFNHND-NAPTM 179
Db      121  GKIIVSCIPPGGSHMLTIAQATLPHVJADVRTLDPIEVPLEDVRNVLFNHNDNQOTM 180
Qy      180  RLVCMLYTLPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNNVEQTKFSPVNLPL 239
Db      181  RLVCMLYTLPLRTGGTG--DSFVWAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNLEL 238
Qy      240  NTLNSRVPSLIKSMVSRDHGMQVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT-----S 294
Db      239  SSLSNRAPLPISSMGISPDNVQVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT-----S 294
Qy      300  GNGYNLTDLGSPYHAFESAPIGFDPDLGECMDHMEASPTQNTGDKVTKQINVKQESA 359
Db      295  NGTVINLTDLGTPFHPEGPAPIGFDPDLGECMDHIN---MTQFGHSSQTYQYDVTTPDT 351
Qy      360  FAPHLGTIQADGLSDVSNTNMIKLGWVSPVSDGHRGVDVPMWIPRYGSTLTTEAAQLAP 419
Db      352  FVPHLGSIQANGIG-----SGNYVGVLSWISPPSHPSGSQVDLWKIPNYGSSITEATHLP 407
Qy      420  PIYPGFGGALVFFMSDPPIAHNGTNGLSVPCITIOEFVTHFVNECAPTRGEAALLHYLDP 479
Db      408  SVYPPGFGELVFFMSKMP---GPGAYNLPCLLPQEIYISHLASQAPTGVGEAALLHYVDP 464
Qy      480  DTHRNLGFEKLYPEGFMTCVNSSGTGPQTLPIGVFVSVWSRFYQKLPVGTAGPA-C 538
Db      465  DTGRNLGFEKAYPDGFLTCVNGASSGQQLPIGVFVSVWSRFYQKLPVGTASSARG 524
Qy      539  RLGIIR 544
Db      525  RLGLRR 530

RESULT 6
AAB49701
ID      AAB49701 standard; protein; 530 AA.
XX
AC      AAB49701;
XX
DT      04-APR-2001 (first entry)
XX
DE      Small round structured virus protein (SEQ ID 2).
XX
KW      Small round structured virus; SRSV; food poisoning.
XX
OS      Small round structured virus.
XX
PN      WO200079280-A1.
XX
PD      28-DEC-2000.
XX
PR      22-JUN-2000; 2000WO-JP004095.
XX
PR      22-JUN-1999; 99JP-00175928.
XX
XX      (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX      (DENK-) DENKA SEIKEN KK.
XX
PI      Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

```

XX WPI; 2001-080848/09.  
 DR N-PSDB; AAF29142.  
 XX  
 XX  
 PT Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.  
 XX  
 XX  
 PS Claim 1; Page 42-45; 84pp; Japanese.  
 XX  
 XX This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX  
 XX Sequence 530 AA;  
 SQ  
 Query Match 66.9%; Score 1938; DB 4; Length 530;  
 Best Local Similarity 67.0%; Pred. No. 4.1e-164;  
 Matches 366; Conservative 63; Mismatches 99; Indels 18; Gaps 7;  
 Qy 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMDPWIMNYY 60  
 Db 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMDPWIMNYY 60  
 Qy 61 VOAPQGEFTISPNNTPGDILFDLQGLPHLNFPLSHLAQMYNGWGNMKVLLAGNAFTA 120  
 Db 61 VOAPQGEFTISPNNTPGDILFDLQGLPHLNFPLSHLAQMYNGWGNMKVLLAGNAFTA 120  
 Qy 121 GKIIISICPPGFAQNTISIAQATFPHVIADRVLEPIEVEPLEDVRNVLPHNND-NAPTM 179  
 Db 121 GKIIISICPPGFSHNLTIAQATLFPFVIADRVLTDLPIEVEPLEDVRNVLPHNNDNQTM 180  
 Qy 180 RLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTPFSPVNLPL 239  
 Db 181 RLVCMLYTPLRTGGGTG--DSFVAVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPL 238  
 Qy 240 NTLNSRVPSLISKSMVSRDHGQVQFQNGRVTLTGQGLQGTPTTSASOLCKIRGSVFHAN 299  
 Db 239 SSLNSRAPLPISGMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT---S 294  
 Qy 300 GGNGYNTLSDGSPYHAFESPAPIGFPLDGLCDHMEASPTTQFNTGDKVINKVQESA 359  
 Db 295 NGTVINLTDLGTFPHFEPGAPIGFPLDGLCDHMEASPTTQFNTGDKVINKVQESA 351  
 Qy 360 FAPHLGTIQADGLSDSVNTNMIKLGWSPVSDGHRGVDVDPWVPIRYGSTLTTEAAQLAP 419  
 Db 352 FVPHLGSIQANGIG---SGNYIGVLSWVSPSPHPSGQVLDLWKIPNYGSSITEATHLAP 407  
 Qy 420 PIYPGGEALVFPMSDFPIAHGNTGLSVPTCTIIOEFVTHFVNEQAPTRGEAALLHYLDP 479  
 Db 408 SVYPPGGEALVFPMSDFPIAHGNTGLSVPTCTIIOEFVTHFVNEQAPTRGEAALLHYLDP 464  
 Qy 480 DTHNLGEFLKYPGGMTCVPSNNGTGPOTLPINGVVFVSWGRFVQKLVGTAGPA-C 538  
 Db 465 DTGRTLEKFXYPGGLTCVPSNNGTGPOTLPINGVVFVSWGRFVQKLVGTAGPA-C 524  
 Qy 539 RLGIIR 544  
 Db 525 RLGIIR 530  
 RESULT 7  
 ID AAR57091  
 XX AAR57091 standard; protein; 530 AA.  
 AC AAR57091;  
 XX  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT

05-OCT-1994 (first entry)  
 Small round virus SRSV/KY/89 capsid protein.  
 pathogen; acute gastroenteritis; food poisoning; seafood contamination;  
 diagnostic assay; human calicivirus; small round virus; SRSV; KY89;  
 Norwalk virus; capsid protein.  
 Small round structured virus.  
 WO9405700-A2.  
 17-MAR-1994.  
 07-SEP-1993; 93WO-US008447.  
 07-SEP-1992; 92US-00941365.  
 (BAYU) BAYLOR COLLEGE MEDICINE.  
 Maseon, DO, 1994, 101125/12.  
 WPI; 1994-101125/12.  
 N-PSDB; AAO56832.  
 DNA from Norwalk and related viruses - used for preparing prods. for use  
 in diagnostic assays, detection and vaccines for Norwalk and related  
 viruses.  
 Example 7; Fig 13a; 156pp; English.  
 The known sequence for Norwalk virus was used to obtain the sequence of  
 other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool  
 from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide  
 cDNA sequence includes part of the polymerase region and the capsid  
 region of the genome; the deduced amino acid sequences are AAR57092 and  
 AAR57091, respectively. Expression of fragments and derivs. of Norwalk-  
 related viruses permits development of diagnostic assays to detect  
 antibodies, antigens, viral genetic material or antivirals. (Updated on  
 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS  
 field.)  
 Query Match 65.6%; Score 1900; DB 2; Length 530;  
 Best Local Similarity 65.7%; Pred. No. 1e-160;  
 Matches 362; Conservative 66; Mismatches 95; Indels 28; Gaps 9;  
 Qy 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMDPWIMNYY 60  
 Db 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMDPWIMNYY 60  
 Qy 61 VOAPQGEFTISPNNTPGDILFDLQGLPHLNFPLSHLAQMYNGWGNMKVLLAGNAFTA 120  
 Db 61 VOAPQGEFTISPNNTPGDILFDLQGLPHLNFPLSHLAQMYNGWGNMKVLLAGNAFTA 120  
 Qy 121 GKIIISICPPGFAQNTISIAQATFPHVIADRVLEPIEVEPLEDVRNVLPHNND-NAPTM 179  
 Db 121 GKIIISICPPGFSHNLTIAQATLFPFVIADRVLTDLPIEVEPLEDVRNVLPHNNDNQTM 180  
 Qy 180 RLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTPFSPVNLPL 239  
 Db 181 RLVCMLYTPLRTGGGTG--DSFVAVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPL 238  
 Qy 240 NTLNSRVPSLISKSMVSRDHGQVQFQNGRVTLTGQGLQGTPTTSASOLCKIRGSVFHAN 299  
 Db 239 SSLNSRAPLPISGMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT---S 294  
 Qy 300 GGNGYNTLSDGSPYHAFESPAPIGFPLDGLCDHMEASPTTQFNTGDKVINKVQESA 354  
 Db 295 NGTVINLTDLGTFPHFEPGAPIGFPLDGLCDHMEASPTTQFNTGDKVINKVQESA 353  
 Qy 355 QKESAFAPHLGTIQADGLSDSVNTNMIKLGWSPVSDGHRGVDVDPWVPIRYGSTLTTEA 414

Db 354 -----PHLGSIQANGIG-----SGNIGVLSWSPSPSHSGSQVDLWKIPNYGSSITEA 402  
Qy 415 AQLAPPYPPGGEAIVFFMSDFPIAHGTNGLSVPTIPQEFVTHFVNEQAPTRGEAALL 474  
Db 403 THLAPSVYSPGGEVLFVFFMSKIP---GPGGDSLPCILLPQGYISHLASEQAPTGVGEPLL 459  
Qy 475 HYLDPDTHRNIGBKLYPEGEMTCVPSNCGTGPOTLPINGVFFVSVWSRFFYQLKPVGTA 534  
Db 460 HYVDPDTHRNIGBKLYPEGEMTCVPSNCGTGPOTLPINGVFFVSVWSRFFYQLKPVGTA 519  
Qy 535 GPA-CRLGIRR 544  
Db 520 STARGELGIRR 530  
RESULT 8  
ID AAB49706 standard; protein; 540 AA.  
XX AAB49706;  
AC AAB49706;  
DT 04-APR-2001 (first entry)  
DE Small round structured virus protein SEQ ID 7.  
DE Small round structured virus; SRSV; food poisoning.  
DE Small round structured virus.  
DE WO200079280-A1.  
DE 28-DEC-2000.  
DE 22-JUN-2000; 2000WO-JP004095.  
DE 22-JUN-1999; 99JP-00175928.  
DE (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
DE (DENK-) DENKA SEIKEN KK.  
DE Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
DE WPI; 2001-080848/09.  
DE N-PSDB; AAF29147.  
DE Kit for the detection and typing of small round-structured virus (SRSV)  
DE strains for investigation of food poisoning outbreaks, contains  
DE antibodies.  
DE Claim 1; Page 54-57; 84pp; Japanese.  
DE This invention relates to a kit for the detection and typing of small  
DE round structured virus (SRSV) strains. The kit contains antibodies  
DE directed against peptides represented in sequences AAB49700 - AAB49710,  
DE which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
DE AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
DE used for detecting and typing strains of SRSV in order to prevent the  
DE spread of infection and to examine the epidemiology of outbreaks  
DE  
DE Sequence 540 AA;  
Query Match 42.3%; Score 1225; DB 4; Length 540;  
Best Local Similarity 46.7%; Pred. No. 2.3e-100; Indels 36; Gaps 14;  
Matches 261; Conservative 89; Mismatches 173;  
Qy 1 MMASKDAPTNWMDGTSGAGQLVPEANTAEPISEMFVAGAAATAAGVNMIDPWIMNNY 60  
Db 1 MKAASNDATPSNDGAAG---LVPEENN-EMAALEPVVGASLAAPVTGTNIIDPWIRNF 56  
Qy 61 VOAPGGEFTISPNTPGDILDLQPLHNPFLSHLAOMYNGVMKVKVLLAGNAFTA 120  
Db 57 VOAPNGEFTVSPRNSPGEILVNLELPELNLARVNYAGGMEVQVNLAGNAFTA 116

Qy 121 GKIIISCIIPGFAAONISIAQATMFPHVIADVRVLEPIEVPLDVRNVLPH-NNDNAPT 179  
Db 117 GKIIIFAAVPPYPPYVENISPSQITNMFPHVIDVRLTEPLVLLPMDVDVRLTEPLHFNQDEPKM 176  
Qy 180 RLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQKTPSPVNLPL 239  
Db 177 RLVAWLYTPLRASG--SGDDVFTVSCRLTRPSPDFTLVLPPTVESKTKPFTLPVLT 234  
Qy 240 NTLNSRVPISLIKMMVSRDQGVQFONGRVTLGGOLGTTPTPSASOLCKIRGSVFHAN 299  
Db 235 GELSNSRPFSLSIDEMVTSPNESIVVQPQNGRVTLDELLGTTQLQACNICIRKVTGQV 294  
Qy 300 GNGY-----NLTELDGSPYHAFES-PAPIGFPDL-GECDWHM-----EASPTTQFNTGD 347  
Db 295 PSEQHMNLLEITNLNGTQFDPTDDVPAPLGVDPDPAGEVFGVLSQRNRCESNPANRAHDAV 354  
Qy 348 VIKQINVKQESAFAPHLGTIQAD--GLSDSVNTNMIAKLGWVSPVSGHGDVDPWVIP 405  
Db 355 V-----ATYSKDYTPKLGVLQIGTWTNTNDVENOPTKFTPIG-LNEVANGHR--FEQWTL 406  
Qy 406 RYGSTLTEAAQLAPPYPPGGEAIVFFMSDFPIAHGTNGLSVPTIPQEFVTHFVNEQA 465  
Db 407 RYSGALTLMNLAPAVAPLFGERLLFRSYVPLKGGFGNPAIDCSVPQEWVQHFYQESA 466  
Qy 466 PTRGEAALLHYLDPDTHRNIGBKLYPEGEMTCVPSNCGTGPOTLPINGVFFVSVWSRFF 525  
Db 467 PSLGDVALVRYVNPDTGRVLFKALHKGGLTV--SSTSTGPPVVVPANGYFKFDSWVNOF 524  
Qy 526 YQLKPVGTAGPACRLGIRR 544  
Db 525 YSLAPMGITGN-----GRRR 538  
RESULT 9  
ID AAB49705 standard; protein; 548 AA.  
XX AAB49705;  
AC AAB49705;  
DT 04-APR-2001 (first entry)  
DE Small round structured virus protein SEQ ID 6.  
DE Small round structured virus; SRSV; food poisoning.  
DE Small round structured virus.  
DE Small round structured virus.  
DE WO200079280-A1.  
DE 28-DEC-2000.  
DE 22-JUN-2000; 2000WO-JP004095.  
DE 22-JUN-1999; 99JP-00175928.  
DE (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
DE (DENK-) DENKA SEIKEN KK.  
DE Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
DE WPI; 2001-080848/09.  
DE N-PSDB; AAF29146.  
DE Kit for the detection and typing of small round-structured virus (SRSV)  
DE strains for investigation of food poisoning outbreaks, contains  
DE antibodies.  
DE Claim 1; Page 52-54; 84pp; Japanese.  
DE This invention relates to a kit for the detection and typing of small  
DE round structured virus (SRSV) strains. The kit contains antibodies  
DE directed against peptides represented in sequences AAB49700 - AAB49710,  
DE which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
DE AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
DE used for detecting and typing strains of SRSV in order to prevent the  
DE spread of infection and to examine the epidemiology of outbreaks  
DE  
DE Sequence 540 AA;  
Query Match 42.3%; Score 1225; DB 4; Length 540;  
Best Local Similarity 46.7%; Pred. No. 2.3e-100; Indels 36; Gaps 14;  
Matches 261; Conservative 89; Mismatches 173;  
Qy 1 MMASKDAPTNWMDGTSGAGQLVPEANTAEPISEMFVAGAAATAAGVNMIDPWIMNNY 60  
Db 1 MKAASNDATPSNDGAAG---LVPEENN-EMAALEPVVGASLAAPVTGTNIIDPWIRNF 56  
Qy 61 VOAPGGEFTISPNTPGDILDLQPLHNPFLSHLAOMYNGVMKVKVLLAGNAFTA 120  
Db 57 VOAPNGEFTVSPRNSPGEILVNLELPELNLARVNYAGGMEVQVNLAGNAFTA 116

CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX  
XX Sequence 548 AA;  
XX

XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA	(DENK-) DENKA SEIKEN KK.
XX	
XX	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
PI	WPI; 2001-080848/09.
XX	N-PSDB; AAF29148.
DR	
XX	
DR	Kit for the detection and typing of small round-structured virus (SRSV)
PT	strains for investigation of food poisoning outbreaks, contains
PT	antibodies.
XX	
XX	Claim 1; Page 57-59; 84pp; Japanese.
PS	
XX	This invention relates to a kit for the detection and typing of small
CC	round structured virus (SRSV) strains. The kit contains antibodies
CC	directed against peptides represented in sequences AAB49700 - AAB49710,
CC	which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC	AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC	used for detecting and typing strains of SRSV in order to prevent the
CC	spread of infection and to examine the epidemiology of outbreaks
XX	
XX	Sequence 535 AA;
QY	Query Match 40.6%; Score 1174.5; DB 4; Length 535;
DB	Best Local Similarity 45.0%; Pred. No. 7.6e-96;
DB	Matches 251; Conservative 87; Mismatches 181; Indels 39; Gaps 13;
QY	1 MMWASDAATNMDGTSGAQLVPEANTAPISMEPVAGAATAAATAGOVNMDPWTNNY 60
DB	1 MMWASNDAAPSNDGAAG---LVPEANN-ETMALEPVAGASIAAPLTGQNNIDPWIRLNF 56
QY	61 VQAPQGEFTISPNNTPGDILFDLQLGPHLNPELASHLAQMYNGVGNMKVLLAGNAFTA 120
DB	57 VQAPNGEFTVSPRNSPGEVLLNLELGPENLPVLAHLRMYNGVAGGVEQVLLAGNAFTA 116
QY	121 GKIIISCIIPGFAAQNISATAQMPFHVADRVLEPIEVLPELDVRNVLFH-NDNAPTM 179
DB	117 GKLVFAAVPHFPLENISPGQITMFPFHVIIDVRTLPEVLLPLDVRNPNFFHYNQNEPRM 176
QY	180 RLVCMLYTPLRASGSSGTDPPFIAGRVLTCPSPDFLFPVPPNVEQTKPFSVNLPL 239
DB	177 RLVMALYTPLRNG--SGDDVTVSCRVLTRSPDFDNLYLPPTLESKTKEFTLLILTI 234
QY	240 NTLNSRVPSPLSIKSMVSRDHGMQVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHAN 299
DB	235 GELTNSRFPVPIDELYTSPNESLVQPPQNGRCALDGELOGTQTLLPTAICSRFRINQKV 294
QY	300 GGNQY---NLTELDSYPH-AFESPAPIGPPD-----LGECDDHWEASPTTQFNTG 346
DB	295 SGENHVWNVQVNTINGTPDPTGDPAPLGTDPFGSLFGVLSQRD-HDNAC-----RSH 348
QY	347 DVIKQINVKQESAFAPHLGTIIQADGLSDYSVNTNIAKLGWSPVSDGHRGVDVDPVPIR 406
DB	349 DAVIATN---SAKFTPKLGAIOIGTWEEDDVHINQPTKF---TPVGLPENEGNQWTLN 402
QY	407 YGSLTEAAQLAPPYPPFGGALVFFPMDFPIANCTNGLSVPCITIQSFVTHFVNEQAP 466
DB	403 YGSCALTNNGLAPPVAPTFFGQIILFFRSHIPLKGVADPVIDCLLPQSWIOHLYQESAP 462
QY	467 TRGEAALLHYLPDTHRNLEGEFKLYPEGFMTCVPNSSGTGPOTLPINGVVFVSWVSRY 526
DB	463 SQSDVALIRFTNPDTCRVLFPEAKLHRSGLVITVA--NTGSRPIVVPANGYFRDPTWVQY 520
QY	527 QLKPVGTAGACRLGIRR 544
DB	521 SLAPMGTGN-----GRRR 533









Db 412 ---TSPTTTPQWRRCSRGSCSHRYVPVVMNRVTWIVLSHKSGFSTSTRKLPQ 468  
Qy 421 IYPEGFGAIVFFMSDFPIAHGTNGLSVPTIPOEFVTHFVNEQAPTRGEALLHYLDPD 480  
Db 469 L-----NLRW-----LIRFINPD 482  
Qy 481 THRLGEFKLYPEGFMTCVPNSSGTGPOTLPINGVFVSVWSRPYQLKPVGT 533  
Db 483 TGRVLFEARLHKQGFITVA--HTGDNPIVMPPNNGYFRFEAWVNOFYSLAPVGT 533

Search completed: June 1, 2004, 13:45:44  
Job time : 48.8188 secs

**THIS PAGE BLANK (USPTO)**



QY 61 VQAPQGEFTTSPNTPGDIILFDLQGLPHLNFSLHQAQYNGVGNMKVLLAGNAFTA 120  
DB 61 VQAPQGEFTTSPNTPGDIILFDLQGLPHLNFSLHQAQYNGVGNMKVLLAGNAFTA 120  
QY 121 GKIIISICPPGFAAONTISIAQATMPFHVIADRVLEPIEVLDPVRLVFNHND-NAPTM 179  
DB 121 GKIIISICPPGFAAONTISIAQATMPFHVIADRVLEPIEVLDPVRLVFNHND-NAPTM 180  
QY 180 RLVCMLYTPRASGSSSGTDPFVIAGRVLCPSDFSLFLVPPNVEOKTKPFSVNLPL 239  
DB 181 RLVCMLYTPRASGSSSGTDPFVIAGRVLCPSDFSLFLVPPNVEOKTKPFSVNLPL 238  
QY 240 NTLNSRVPSLISKMSVSRDHQGVQFONGRVTLTGQLOQTTPTSASQCKIRGSVFHAN 299  
DB 239 SSLNSRAPLPISSMGTSIPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGT-----S 294  
QY 300 GGNLYNLTELDPGSPVHAPESAPITGFPDLGCDWHMEASPTQNTGDVVIKQINVKQESA 359  
DB 295 NGTVNLTELDPGSPVHAPESAPITGFPDLGCDWHMEASPTQNTGDVVIKQINVKQESA 351  
QY 360 FAPHLGTTQADGLSDVSN-TNMI-----AKLGVSPVSD-----GHRGDVDPWVI 419  
DB 352 FVPHLGSIQANGIG-----SGNVGVLSWISPPSHSGSOVDLWKIPNVGSSITEATHLAP 407  
QY 420 PIYPPGFGAEIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGAEALLHYLDP 479  
DB 408 SVYPPGFGAEIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGAEALLHYLDP 464  
QY 480 DTHRNGLBEFLYPEGFMTCVPSNCGTGTPLINGVVFVSVWSRFFYOLKPVGTAGPA-C 538  
DB 465 DTGRNLGEPKAYPDGFLTCVPSNCGTGTGTPLINGVVFVSVWSRFFYOLKPVGTAGPA-C 524  
QY 539 RLGIR 544  
DB 525 RLGIR 530

RESULT 2  
US-09-590-020-7  
; Sequence 7, Application US/09590020  
; Patent No. 635246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Aivars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4-1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; PRIOR FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-7  
Query Match 9.8%; Score 285; DB 4; Length 626;  
Best Local Similarity 23.8%; Pred. No. 6.9e-20;  
Matches 144; Conservative 87; Mismatches 194; Indels 180; Gaps 32;  
QY 13 DGTSGAGQLVPEANT-AEPTSMPEVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI- 70  
DB 82 DGSITA-----PEQGTWVGVIAPESQAQMTAADATKGSVDSEW-----EAFFSFH 128  
QY 71 -----SPNTPGDIILFDLQGLPHLNFSLHQAQYNGVGNMKVLLAGNAFTAAGKIII 125  
DB 129 TSVNWSSTSETQKILFKQSLGLPLNPNLYSLAKLYVWWSGSIEVRFPSISGSGVFGKLA 188

QY 126 SCIPPGF-AAQNIISIAQATMPFHVIADRVLEPIEVLDPVRLVFNHNDNAPTMRLVCM 184  
DB 189 IVPPGVDVOSTSMQO---YPHVLFDAQVEPVIFCLPDLRSTLYHLMSDITDTTSLVIM 245  
QY 185 LY-----TFLRASGSSSGTDPFVIAGRVLCPSDFSLFLVPPNVEOKTKPFSVNLPLN 240  
DB 246 VYNDLINYANDTSSGC---IVT---VETKPGDFKPHLLKPPG-----S 285  
QY 241 TFLNSRVPS-LI---KSMVSRDH-QGMVQFONGRVTLTGQLOQTTPTSASQCKIRGSV 295  
DB 286 MLTHGVSFSDLIKPSKSSSLWIGNRHWSIDTF-----LIRFV 322  
QY 296 PHANGGVN-----LTLDGSPYHA-----PESPA-PIGFPDL-----G 329  
DB 323 FOANRHFENQTAGWSTPRPISVTITEQNGAKLIGVATDYIVPGIDGWDPTTIPG 382  
QY 330 EC-----DWHM-----EASPTQNTGDVVIK-QINVK-----QESAPA 361  
DB 383 ELIPAGDYAITNGTGNDDITATGYDTADIIKNNTFRGMVTCGSLQORAWGDKKISNTAP- 441  
QY 362 PHLGTITQADGLSDVSN-TNMI-----AKLGVSPVSD-----GHRGDVDPWVI 404  
DB 442 --ITATLTDGNNKINPCNTIDQSKIYVQDAHVKKQAQTSDDTLALLGYTGIGEQAI- 498  
QY 405 PRYGSTLTEAAQLA--PPIYPPGFGAEIVFFMSDFPIAHGTNGLSVPCCTIPOEFVTHFVN 462  
DB 499 ---GSDRDRVRISTLPETGARG-GNHPIFYKNSIKLGYVIRSIDV-----FNS 543  
QY 463 EQAPTRGAEALLHYLDP-----DTHRNGLBEFLYPEGFMTCVPSNCGTGTPLI 512  
DB 544 QILHTSRQLSNLHYLLPPDSFVYRIIDNSGWSFIDIGSDGF--SFVGSVFGKLEPDL 601  
QY 513 NGVVF 517  
DB 602 SASYM 606

RESULT 3  
US-09-617-594A-2  
; Sequence 2, Application US/09617594A  
; Patent No. 6541458  
; GENERAL INFORMATION:  
; APPLICANT: Augonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V. V  
; FILE REFERENCE: 454313-3151.1  
; CURRENT APPLICATION NUMBER: US/09/617,594A  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR FILING DATE: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 2  
; LENGTH: 669  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-617-594A-2  
Query Match 9.4%; Score 272.5; DB 4; Length 669;  
Best Local Similarity 25.9%; Pred. No. 1.4e-18;  
Matches 106; Conservative 54; Mismatches 134; Indels 115; Gaps 19;  
QY 13 DGTSGAGQLVPEANT-AEPTSMPEVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI- 70  
DB 126 DGDSSI---TTPEQGTWVGVIAPESQAQMTAADATKGSVDSEW-----ESFFSFH 174  
QY 71 -----SPNTPGDIILFDLQGLPHLNFSLHQAQYNGVGNMKVLLAGNAFTAAGKIII 125  
DB 175 TSVNWSSTSETQKILFKQSLGLPLNPNLYSLAKLYVWWSGSVDVRFPSISGSGVFGKLA 234

126	Qy	SCI	PGP	-	AAQ	ISIAQ	ATMP	PHV	IA	VR	LE	PI	EP	LE	VR	VL	F	H	N	D	N	A	P	T	R	L	V	C	M	184	
235	Db	I	VP	PG	V	D	P	V	O	S	T	M	L	O	---	Y	P	H	L	F	D	A	R	Q	V	E	P	I	F	S	291
185	Qy	L	Y	---	-	T	P	L	R	A	S	G	S	S	G	T	D	P	F	V	I	A	G	R	V	L	T	C	P	S	240
292	Db	V	Y	N	D	L	I	N	P	A	N	S	S	G	---	I	V	T	---	V	E	T	K	G	P	D	F	K	H	L	331
241	Qy	T	L	S	N	R	V	P	S	---	L	I	K	S	M	V	---	S	R	D	H	G	Q	M	V	Q	F	O	N	G	295
332	Db	M	L	T	H	G	S	I	P	S	D	L	I	P	K	S	S	S	L	I	G	N	Y	M	S	D	I	T	D	---	368
296	Qy	P	H	A	G	G	G	G	N	---	---	---	---	---	---	---	---	L	T	E	L	D	G	S	P	Y	H	---	F	330	
369	Db	F	O	A	N	R	H	F	D	F	N	O	E	T	A	G	S	T	P	R	P	I	T	I	S	E	N	G	S	K	428
331	Qy	C	D	H	M	E	A	S	P	T	T	F	N	G	D	V	I	K	I	N	V	K	E	S	A	F	A	P	H	L	379
429	Db	---	E	L	T	P	A	G	D	S	I	---	---	---	---	---	---	T	N	G	S	G	N	D	I	A	T	A	N	461	

RESULT 4  
US-09-590-020-4  
; Sequence 4, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Aivars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALCIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-4

[illegible]

```

Db      400 ITTAAG--YDAAEATVNTNFKSMYICGSLQAMGDKKISNTAFITTAVRKGNSEPSNT 457
          : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Qy      346 GDVIKQINVKQESAFAPHLGTIQADGLSDSVSNVTNMIAKLGVSPVSDGHRGVDVP--- 401
          : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db      458 IDMTKLV-VYQDA-----HVG-----REVQTSIDTALLGYTGIGEEAIGSDRDKVVRI 505
          : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Qy      402 WVIPRYGSTLTTEAAQLAPPYPPGFGEAIVFMSDFPIAKGTNGLSVPCCTIPOEFVTHFV 461
          : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db      506 SVLPETGAR-----CGNHPIFYKNSIKLGVYIRSIDV-----FN 539
          : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Qy      462 NEQAPTRGEAALHHVLDP-----DTHRNLGEPFLYPGGFMTCVPNSSGTGPQTLP 511
          : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |
Db      540 SQILHTRSQSLNNYLPPDSFAVYRIILDSNGSWFDIGDITDGF-----SFVGSVNLP 592
          : | | | | | : | | | | | : | | | | | : | | | | | : | | | | |

RESULT 5
US-09-617-594A-4
; Sequence 4, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR REC
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-4
```

Query Match	9.4%	Score 272;	DB 4;	Length 568;
Best Local Similarity	26.7%;	Prod. No. 1.6e-18;		
Matches 107;	Conservative 46;	Mismatches 134;	Indels 114;	Gaps 201
Qy	23	PEANT-----ASP-ISMEPVAGATAAAATAGOVNMDPWINVYVQAPQCEFTI-----	70	
Db	133	PEQTLVGGVIAEPNAQMSAVADVATGKSDSE-----W-----EAFPSHTSVN	177	
Qy	71	SPNPTGDI LFDLQPHLPNF LSHLAQMYNGVGNMKVKVLLAGNAFTAGKIISCIP	129	
Db	178	WSTSETQGLFKFQSLGFLNPNLTHLAKLYVAMVSGSIEVRSISGSGVFGKLAALIVP	237	
Qy	130	PGF-AAQNI SI AQTMEPHVIADVRLIEVPLEVDYRNVL FHNNDNAPT MRLVCMLY--	186	
Db	238	PGIDPVOSTSMLG---YPHVLFDQARQVEPVIFFTIPDLRNSLYHLMSD TDTTSLVIMIVND	294	
Qy	187	--TPLRASGSSSGTDPPVIAAGRVLTCPSPDPFSFLVLPVPPNVEQTKPFSVENPLNTLSN	244	
Db	295	LINPYANDSNSGC---IVT--VETKGPDPFKFHLKPPG-----SMJTH	334	
Qy	245	SRVPS-LI-----KSMVSRDH-GQMVQFQNGRVLTDGLQGTTPTSASQLCKIRGSVFHAN	299	
Db	335	GSIPSDLI PKSSSLIGNRHMSDITDP-----VIKPFVFOQN	371	
Qy	300	GGNGYN-----LTELDDGSPYH---APESPAPIGFPDLGECDMHMEASP	339	
Db	372	RHDFNQETAGWSTPRFPITITVSEKGSGLGIGVATDSIVP-GIPD-----GWPDTTIP	426	
Qy	340	TTQNTGDVTKIQNVKQESAPAPHLGTIQADGLSDSVNTN	380	
Db	427	EKLTPAGDYAITNGNNDITTA-----ADYDGASIIKNNTN	462	

```

Query Match      9.4%; Score 272; DB 4; Length 668;
Best Local Similarity 26.7%; Pred. No. 1.6e-18;
Matches 107; Conservative 46; Mismatches 134; Indels 114; Gaps 20;

Qy      23 PEANT-----ASP-ISMEPVAGAAATAAATAGOVNMIDPMIMNYVQAPOGEPTI----- 70
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      71 -SPNPTGDIPLDQLQGLPHLPFLSHLAQMYNGWGMKVLLAGNAFTAGKIIISCIP 129
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      178 WSTSETQKILPKQSLGSLPNLPYLTHLAKLYAVWMSGIEVRFSGISGVFGGKLAATVWP 237
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      130 PGF-AAQNTISIAQTWPFPHVIADVRLIEVLEPVEDVRNVLFHNNNDNAPMTRLVCMLY-- 186
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      238 PGIDPVQSTSMLO---YPHVLFDARQVEPVITFPDLRNSLYHLMSTDTTSLVIMLYND 294
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      187 --TPLRASGSSGTDPFVFIAGRVLTCPSPDFSEFLFVPPNVVEQTKPFSPVNPFLPLTSLN 244
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      295 LINPYANDSNSGC---IVT--VETKPGDPKFHLKPPG-----SMLTH 334
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      245 SRVPS-LI-----KMMVSRDH-GQWVQFQNGRVTLDGQLGTTTTSASOLCKIRGSVPFAN 299
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      335 GSIPSDLIPKSSSLWGNHMSDITDF-----VIRPFVFQAN 371
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      300 GGNGYN-----LTLDGSPYH---AFESPAPIGFPDLGECDOHMEASP 339
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      372 RHDFPQETAGMSTPRPRPITIIVSEKSGKLGIIVATDSIVP-GIPD-----GWPDTTIP 426
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      340 TTQFNTGDIKVINVKQESAFAPHLGTIOADGLSDSVNTN 380
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy      427 EKLTTPAGDYAITNGGNNDITTA-----ADYDGASIIKNTN 462
Db      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 6
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Meas, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590.020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-2

Query Match 9.3%; Score 269; DB 4; Length 623;
Best Local Similarity 29.8%; Pred. No. 2.9e-18;
Matches 76; Conservative 41; Mismatches 94; Indels 44; Gaps 9;

QY 7 DAPT-NMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAAGQVNMIDPIMNNYVQAPQ 65
DB 72 DLPLFRLEGDDGSITTPGCTMVGVIAEPSSAQMSAADMATGKSVDSW-----E 122
QY 66 GEFTI-----SPNPTPGDILFDLQGLPHLNPFLSHLAQMYNGVGNMKVLLAGNAFT 119
DB 123 AFPSFHTSVNWSSTETQKILFKQSLGPLNLPYLHLAKLYVALAGSVEVRFSISGSGVF 182
QY 120 AGKIIISCIPPGF-AAQNISIAQTMFPHVIAQVLEPIEVLVDVRLVFNHNDNAPT 178
DB 183 GKGKLAIVVPPGIEPVQSTMLQ---YPHVLFDAQVEVIFAIPDLRSNLHLMSDITD 239
QY 179 MRLVCMXY-----TPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFSV 234
DB 240 TSLVIMVYNDLINPYANDTNSSGC---IVT--VETKPGDPFKHLLKPPG----- 284
QY 235 PNLPLNTLSNSRVS-LI---KSMVSRDHQMVQ-----FQ-NGRVTLDGLOQ-- 279
DB 285 -----SMLTHGSPVPSDLIPKSSSLWIGNRHSWSDITDFIRPFVQANRHFDFNOETAGWS 339
QY 280 -----TTPTSASQLCKI-----RGSVFHANGGNGYN 305
DB 340 TPRFRPITTVSESNMKGIGVATDYIVPGIDGWPDTTPEQLTPAGIYSITASNGTD 399
QY 306 LTELDGPYHAFESPA-PIGFDPDLGEC-----WHMEASPTTQF-----NT 345
DB 400 ITTAAG--YDAETIVNTTFKSMVICGSLQRAWDDKISNTAFITTAVRKGNSTEPSNT 457
QY 346 GDVIKQINVKQSSAFAPHLGTTIQAQGLSDVSNTNMIAKLGWSPVSDGHRGDVDPWV-I 404
DB 458 IDMTKLIV-VYQDA---HVG-----EVQTSIDITLALLGYTGIGEEAIGSDRDKVRI 505
QY 405 PRVGSSTLTAQAQLAPIYPPGGEAIVFPFMSDFPIAHGTNGLSVFCCTIPQEPVTHFVNEQ 464
DB 506 SVLGETGARG-----GNHPIFYKNSIKLGYVIRSIDV-----FNSQI 542
QY 465 APTRGEAALHY-LDPDTHR-----NLGEFKLYPEGFMTCVPSNSGSGTGPQLP 511
DB 543 LHSSRQLSLNLYLPPDPSFAVIRLDSNGSWFDIGIDTFEPDGF-----SFVGVSNLP 595

RESULT 7
US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Meas, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590.020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-6

Query Match 9.1%; Score 262.5; DB 4; Length 622;
Best Local Similarity 22.3%; Pred. No. 1.3e-17;

```

```

Matches 134; Conservative 82; Mismatches 213; Indels 171; Gaps 28;

QY 7 DAPT-NMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAAGQVNMIDPIMNNYVQAPQ 65
DB 72 DLPLFRLEGDDGSITTPGCTMVGVIAEPSSAQMSAADMATGKSVDSW-----E 122
QY 66 GEFTI-----SPNPTPGDILFDLQGLPHLNPFLSHLAQMYNGVGNMKVLLAGNAFT 119
DB 123 AFPSFHTSVNWSSTETQKILFKQSLGPLNLPYLHLAKLYVALAGSVEVRFSISGSGVF 182
QY 120 AGKIIISCIPPGF-AAQNISIAQTMFPHVIAQVLEPIEVLVDVRLVFNHNDNAPT 178
DB 183 GKGKLAIVVPPGIEPVQSTMLQ---YPHVLFDAQVEVIFAIPDLRSNLHLMSDITD 239
QY 179 MRLVCMXY-----TPLRASGSSGTDPPFVIAGRVLTCPSPDFSLFVPPNVEQTKPFSV 234
DB 240 TSLVIMVYNDLINPYANDTNSSGC---IVT--VETKPGDPFKHLLKPPG----- 284
QY 235 PNLPLNTLSNSRVS-LI---KSMVSRDHQMVQ-----FQ-NGRVTLDGLOQ-- 279
DB 285 -----SMLTHGSPVPSDLIPKSSSLWIGNRHSWSDITDFIRPFVQANRHFDFNOETAGWS 339
QY 280 -----TTPTSASQLCKI-----RGSVFHANGGNGYN 305
DB 340 TPRFRPITTVSESNMKGIGVATDYIVPGIDGWPDTTPEQLTPAGIYSITASNGTD 399
QY 306 LTELDGPYHAFESPA-PIGFDPDLGEC-----WHMEASPTTQF-----NT 345
DB 400 ITTAAG--YDAETIVNTTFKSMVICGSLQRAWDDKISNTAFITTAVRKGNSTEPSNT 457
QY 346 GDVIKQINVKQSSAFAPHLGTTIQAQGLSDVSNTNMIAKLGWSPVSDGHRGDVDPWV-I 404
DB 458 IDMTKLIV-VYQDA---HVG-----EVQTSIDITLALLGYTGIGEEAIGSDRDKVRI 505
QY 405 PRVGSSTLTAQAQLAPIYPPGGEAIVFPFMSDFPIAHGTNGLSVFCCTIPQEPVTHFVNEQ 464
DB 506 SVLGETGARG-----GNHPIFYKNSIKLGYVIRSIDV-----FNSQI 542
QY 465 APTRGEAALHY-LDPDTHR-----NLGEFKLYPEGFMTCVPSNSGSGTGPQLP 511
DB 543 LHSSRQLSLNLYLPPDPSFAVIRLDSNGSWFDIGIDTFEPDGF-----SFVGVSNLP 595

RESULT 8
US-07-852-260-2
; Sequence 2, Application US/07852260
; Patent No. 5525715
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 30 Rockefeller Plaza
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/852.260
; FILING DATE: 19920619
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678

```

```

; ; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US
; ;
; ; TELECOMMUNICATION INFORMATION:
; ; TELEPHONE: (212) 977-9550
; ; TELEFAX: (212) 664-0525
; ; TELEX: 422523 COOP UI
; ; INFORMATION FOR SEQ ID NO: 2:
; ; SEQUENCE CHARACTERISTICS:
; ; LENGTH: 2206 amino acids
; ; TYPE: AMINO ACID
; ; TOPOLOGY: linear
; ; MOLECULE TYPE: protein
; ; US-07-852-260-2

```

Query Match 5.1%; Score 149; DB 1; Length 2206;  
Best Local Similarity 20.3%; Pred. NO. 3.8e-05;  
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

[illegible]

```

RESULT 9
US-08-461-503--2
; Sequence 2, Application US/08461503
; Patent No. 5834302
; GENERAL INFORMATION:
; APPLICANT: Racanietello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES
; TITLE OF INVENTION: FROM CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

```

;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,503
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-503-2

Query Match      5.1%; Score 149; DB 2; Length 2206;
Best Local Similarity 20.3%; Pred. No. 3.8e-05;
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

Qy 15 TSAGAGLVPEANTAEFISHEPVAGATAAATAGQV-NM-----IDPWIMNNYVQAPQG-- 66
Db 347 TPGSNQYLTSNDHQSPCAI-PEFDVTPPIDPGEVRKMMELAEIDTWIPLNLESTKRNTM 405
Qy 67 ---EFTISPNNTPGDILFDLQGLPHLNPEFLSH-----LAQMYNGWGNMKVKVLLAGNAP 118
Db 406 DMYRVTLSDSADLSQPILCLSLSPAPDRLSHTMLGEVLNYYTHWAGSLKFTFLFCGMM 465
Qy 119 TAGKIITSCIPPGFAAQNTSIAQATMFPHPVIADVRVLEPIEVPLEDVRNVLFH--NNDNA 176
Db 466 ATGKILVAVAPG-AQPPSTRKEMLGTHVINDLGLQSSCTMVVPVINSVYRQTQDSP 524
Qy 177 PTWRLVCMLY-----TPLRASGSSSTGDPFVIAGRVLTCPSPDFLFL-----V 221
Db 525 TEGGYISMFQYTRIVVPLSTPKSMS-----MLGFVSAC--NDFSVRLLADTTTHISQSA 576
Qy 222 PNVVEQKTK-----PFSVNP.L-----LNTLSNSRVSLIKSM 255
Db 577 PQGIEDLTSEVAGALTSLPKQDLSLPTKASGPAHSKEVPALTAVETGATNPLAPSDT 636
Qy 256 VSRDHGQWQ-----FQNGRVTLDGLOGTTPTSASQLCKIRGSVFH----- 297
Db 637 VQTRH--VVQRRSRSESTIESPFARGACVAIEVDNEQPTTAAKLFAMWRITYKDTVOL 694
Qy 298 -----ANGSGYNLTLDGSPYHAFESPAPIGFPDLGECDW 333
Db 695 RRKLEFPYTSRDFMEFTVVVTANFTNANGHALNQVYQIMYIPPGAPTPKSWDY---TW 751
Qy 334 HMEASTTQNTGDKVIQINVKQESAFAPHLGTIQIA-----DGLSDVSVNTNMIKLG 386
Db 752 QTSNPSIFITYGAAPARISV-----PYVGLANAYSHFYDGFVAKVPLKTDANDQIG 802

RESULT 10
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York

```

COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release 1.30  
CURRENT APPLICATION DATA: US/08/465,250  
APPLICATION NUMBER: US/08/465,250  
FILING DATE: 6-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 36607-E-PCT-US  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2206 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-465-250-2

Query Match 5.1%; Score 149; DB 3; Length 2206;  
Best Local Similarity 20.3%; Pred. No. 3.8e-05;  
Matches 97; Conservative 57; Mismatches 196; Indels 128; Gaps 19;

15 TSGAGQLVPEANTAEFISMEPVAGATAAATAGOV-NM-----IDPWNNVVOAQQ-- 66  
347 TPGSNQYLSDNHQSPCAI-PEFDVTPPIDGKVMMLAEIDTMIPLNESTKRTM 405  
67 ----EFTISPNPTGDLFDLQGLPHLNPELH-----LAQVNGVGVNKKVLLAGNAF 118  
406 DMVRVTLSDSADLSQPICLSLSPAPDRLSHMTLMGEVLNYTHWAGSLKFTFLFCGSMW 465  
119 TAGKIIICIPPGFAAQNISIAQATMFPHVIADRVLEPIEVLEDRVNLPH--NNNA 176  
466 ATKILVAVAPFG-AQPTSRKEAMLGTHVINDGLQSSCTMVVFWISNVTYRQTQDSF 524  
177 PTWRLVCMLY-----TPLRAGSSSGTDFVIAGRVLTCPSPDFSLFL-----V 221  
525 TEGGYISMFTQTRIVVPLSTPKSMS-----MLGFVSAC--NDFSVLLRDTHISQSAL 576  
222 PNVEQKTK-----PFSVPNLP-----LNTLSNRVPSLIKSNM 255  
577 PQGLDUTSEVAQALTLSPKQODSLPDTKASGPAHSKEVPALTAVETGATNPLAPSDT 636  
256 VSRDHGQMVQ-----FQNGRVTLDDGLQGTTPTSASOLCKIRGSVFH----- 297  
637 VQTRH--VVQRRSRSESTIESFFARGACVAIIEVDNEQTPTRAQKLFAMWRITYKDTVOL 694  
298 -----ANGGVNLTDLGSPYHAFESPAPIGFDPDLGECDW 333  
695 RRKLEFTYSRDFMEFTVVVTANFTNNGHALNQVQIMYIPPGAPTPKSWDY--TW 751  
334 HMEASPTQFTGTVIKQINVKQESAFAPHLGTIOA-----DGLSDVSVNTNMIKLG 386  
752 QTSNPSIFYYTGAAPARISV-----FVYGLANAYSHFYDGFAPVPLKTDANDQIG 802

RESULT 11  
US-09-252-991A-26695  
; Sequence 26695, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136

CURRENT APPLICATION NUMBER: US/09/252,991A  
CURRENT FILING DATE: 1999-02-18  
PRIOR APPLICATION NUMBER: US 60/074,788  
PRIOR FILING DATE: 1998-02-18  
PRIOR APPLICATION NUMBER: US 60/094,190  
PRIOR FILING DATE: 1998-07-27  
NUMBER OF SEQ ID NOS: 33142  
SEQ ID NO 26695  
LENGTH: 615  
TYPE: PRT  
ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-26695

Query Match 3.7%; Score 106; DB 4; Length 615;  
Best Local Similarity 22.0%; Pred. No. 0.11;  
Matches 107; Conservative 56; Mismatches 181; Indels 142; Gaps 24;

4 ASKDAPTNMDGTSGAGQLVPEANTAE-----ISMEPVAGATAAATAGOVNMID 53  
133 SSSTAP-----RPGPGVLPVATPSSPPALGSPGVIPVEPPNSPPAPKSPVPGGVVPE 187  
54 PWIMNYVOAQQ-----GEFTISPNNT-----PGDILFDLQGLPHLPFLSLHAQMYNG 102  
188 P---PNSPPAPKSPVPGGVVPEPPNSPPAPKSPVPGGV-----PVAPSSSTAPR--- 234  
103 WVGNNMKVLLAGNAFTAGKIIISCIPPGFAAQNISIAQATMFPHVIADRVLEPIEVPL 162  
235 -PGPGVLPVATPNSSTAPRSILGRVLP-VATPNSPPAPKSPVPGGV-----PVEPP- 284  
163 EDVRNVLFNNDNAPTMRLVCMLYTPLRASGS-----SSGTDPFVIAG 205  
285 -----NSPPAPKSPVPGGVVPEPPNSPPAPKSPVPGGVVAPVAPNSSTAPRSILG 333  
206 RVLTCPSPDFSLFLVPPNVEQKTKPFVSN--LPLNTLSNRVPSLIKSNMVRDHOQM 263  
334 RVLPAVATPS-----SPPALG-----SVPGGVIPVESPSNSTAP---RSIL----- 370  
264 VQFQNGRVTLDDGLQGTTPTSASOLCKIRGSVPHANGN-----GYNLTELDGSP 313  
371 -----GRV-----LPVATPSSPPALGSPGVIPVEPPNSPPAPKSPVPGGVVPEPPNSP 420  
314 YHAFESPAPIGFDPDLGCDWHMEASPTQNTGDVIKQINVKQESAFAPHLGTIOADGLS 373  
421 -----PAPKSPV--GEV--VPVAPNSSTAPRSILGRVLPVATPSSPPALGSPVPGGVIP 470  
374 DVSVNTNMIK--LGWVSPV-----SDGHRGDVDPWIPRYGSTLTAAQLAP---PIYP 423  
471 VESPSNSTAPRSILGRVLPVATPSSPPALGSPGVIPVESPSNSTAPRLDPEGVVVPAP 530  
424 PGFGEA 429  
531 PSSSTA 536

RESULT 12  
US-09-423-890-2  
; Sequence 2, Application US/09423890  
; Patent No. 6312934  
; GENERAL INFORMATION:  
; APPLICANT: CADUS PHARMACEUTICAL CORPORATION  
; TITLE OF INVENTION: HUMAN MEK PROTEIN AND NUCLEIC ACID MOLECULES  
; TITLE OF INVENTION: AND USES THEREFOR  
; FILE REFERENCE: CPI-085CPPC  
; CURRENT APPLICATION NUMBER: US/09/423,890  
; CURRENT FILING DATE: 2000-03-06  
; PRIOR APPLICATION NUMBER: USN 60/078,153  
; PRIOR FILING DATE: 1998-03-16  
; PRIOR APPLICATION NUMBER: USN 60/099,165  
; PRIOR FILING DATE: 1998-09-04  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 1302





Db 867 EFYPHIVSTVDSQAEPEVIRYKLLSLTALQSIDNSHVMGKLSRRIY----- 916  
Qy 197 GTDFPVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSPVNLPLNTLSNRVPSLIKSMVMV 256  
Db 917 -----LSARWVTAVPAVFSKLVTM-----LNASGTHFTMRRLMA 954  
Qy 257 SRDHQVMQV-FQNG-RVTLDG---QLQTTPTSQCLCKIRGSVFHANGNGVNLTELD- 310  
Db 955 IADEVEIAEVIQLGVEDVDGHDLSQAVPTSCLENSLSLEHTVHREKTKGKLSATRLSA 1014  
Qy 311 -----GSPYHAFESPAPIGFDPDLGECDDHMEAS 338  
Db 1015 SSEDISDLRAGVSVGLPSTTTEQKPAVQTKGRPHSQCLNSPLSHAQL-----MPPA 1068  
Qy 339 PTTQNTGDNVQINVKQESAFAP 362  
Db 1069 PSAPCSAPSVPDISKRRPQAFVP 1092

RESULT 15  
US-09-091-219-24  
; Sequence 24, Application US/09091219  
; Patent No. 6171592  
; GENERAL INFORMATION:  
; APPLICANT: STUDDERT, Michael J.  
; APPLICANT: CRABB, Brendan S.  
; APPLICANT: FENG, Li  
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS  
; FILE REFERENCE: 040268/0151  
; CURRENT APPLICATION NUMBER: US/09/091,219  
; CURRENT FILING DATE: 1998-10-05  
; EARLIER APPLICATION NUMBER: PCT/AU96/00815  
; EARLIER FILING DATE: 1996-12-18  
; EARLIER APPLICATION NUMBER: AU PN7201  
; EARLIER FILING DATE: 1995-12-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 2318  
; TYPE: PRT  
; ORGANISM: Foot-and-mouth disease virus  
US-09-091-219-24

Query Match 3.7%; Score 106; DB 3; Length 2318;  
Best Local Similarity 22.3%; Pred. No. 0.96;  
Matches 63; Conservative 31; Mismatches 113; Indels 76; Gaps 13;  
Qy 5 SKDPTNMDCTSGAGQLVPEANTAEPISME----PVAGATAAATAGQVNMIDPWIMNYY 60  
Db 270 SKLASSAFSLFGA--LLADKTEETLLEDRILTTTRNGHTTSTTQSSVGV-----TY 320  
Qy 61 VQAPQGEFTTSPNNTPGDI-----LFDLQLGP-----HL-----NPF 92  
Db 321 GYATAEDFVGPNSTSGLETRVQAEFPFKTHLFDWTSDFGRCHELLELFDTHKGVYGS 380  
Qy 93 LSHLAQMYNGVGNMKYKLLAGNAFTAGKIIISCIPPGFAQNIISIAQATMPHPHVIADV 152  
Db 381 TDSYAYMRNGW----DVEVTAVGNQFNGGCLLVAMWPELYSIQKRELYQLTLPHQFINP 436  
Qy 153 R--VLEPIEVPLEDRVNV-----LFHNDNAPTWRLLVCMLY--TPLRA 191  
Db 437 RTNNTAHTVFPVGVNRYDQYKHKPWTLVVMVAPLTNTAGAPQIKVYANIAPTNVHV 496  
Qy 192 SG-----SSSGTDPFVIA-----GRVLTCP-----SPDFSFLFLVPPN 224  
Db 497 AGEFPFSKEGIFPVACSDGYGGLVTTDPKTAADPVYGVNFPNPN 539

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.8398 Seconds  
(without alignments)  
4368.312 Million cell updates/sec

Title: US-09-926-799-1  
Perfect score: 2896  
Sequence: 1 MMASKDAPTNMDGTSGAQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071 .

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1950	67.3	530	14	US-10-314-739-3
2	276	9.5	547	12	US-10-670-695-4
3	276	9.5	671	12	US-10-670-695-2
4	272.5	9.4	669	14	US-10-209-507-2
5	272	9.4	668	14	US-10-209-507-4
6	122.5	4.2	6310	12	US-10-282-122A-67793
7	122	4.2	913	15	US-10-369-493-2294
8	122	4.2	3263	12	US-10-282-122A-77663
9	114.5	4.0	1765	12	US-10-282-122A-48055
10	113	3.9	724	14	US-10-156-761-9262
11	113	3.9	1555	14	US-10-128-714-3298
12	113	3.9	1832	14	US-10-128-714-8298
13	112.5	3.9	2914	15	US-10-093-463-82
14	112	3.9	3930	12	US-10-282-122A-46817
15	111	3.8	1046	14	US-10-224-999A-3480

16	108.5	3.7	1194	12	US-10-282-122A-46163
17	108	3.7	5935	14	US-10-243-243A-8
18	107.5	3.7	585	15	US-10-369-493-15469
19	107.5	3.7	2703	12	US-10-282-122A-66108
20	107	3.7	5877	14	US-10-142-515-11
21	106	3.7	507	12	US-10-424-599-185988
22	106	3.7	510	12	US-10-425-114-72837
23	106	3.7	1302	13	US-10-000-864-2
24	106	3.7	1493	9	US-09-858-754-3
25	106	3.7	1493	13	US-10-000-864-8
26	105.5	3.6	818	14	US-10-174-677-86
27	105.5	3.6	932	14	US-10-174-677-87
28	104.5	3.6	946	11	US-09-764-875-638
29	104	3.6	435	15	US-10-609-775-24
30	104	3.6	800	15	US-10-108-260A-2490
31	104	3.6	1052	12	US-09-905-211-2
32	103.5	3.6	1170	13	US-10-135-687-2
33	103.5	3.6	1210	9	US-09-860-352A-2
34	103.5	3.6	1210	15	US-10-410-764-89
35	103.5	3.6	1210	15	US-10-258-106-1
36	102.5	3.5	869	12	US-10-087-192-1683
37	102.5	3.5	1210	12	US-09-963-131-197
38	102	3.5	1048	14	US-10-223-538-10
39	101.5	3.5	1021	16	US-10-443-101-2
40	101	3.5	616	14	US-10-156-761-10270
41	100.5	3.5	1070	9	US-09-735-367B-6
42	100	3.5	544	14	US-10-174-693-349
43	100	3.5	828	9	US-09-738-626-5038
44	100	3.5	1283	15	US-10-369-493-22781
45	100	3.5	1589	15	US-10-369-493-5611

ALIGNMENTS

RESULT 1  
US-10-314-739-3  
Sequence 3, Appli  
Publication No. US20030129588A1  
GENERAL INFORMATION:  
APPLICANT: Estes, Mary K  
Jiang, Xi  
Graham, David Y  
TITLE OF INVENTION: Methods and Reagents to Detect and  
Characterize No. US20030129588A1walk and Related Viruses  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski L.L.P.  
STREET: 801 Pennsylvania Ave., N.W.  
CITY: Washington, D.C.  
STATE: <Unknown>  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/314,739  
FILING DATE: 09-Dec-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,049  
FILING DATE: June 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Davis, Peter  
REGISTRATION NUMBER: 36,119  
REFERENCE/DOCKET NUMBER: 311.023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-662-0200  
TELEFAX: 202-662-4643  
TELEX: <Unknown>

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 530 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match      67.3%; Score 1950; DB 14; Length 530;
Best Local Similarity 67.0%; Pred. No. 6.1e-180; Indels 18; Gaps 7;
Matches 366; Conservative 66; Mismatches 96;

QY 1 MMASKDPTNMDGTSAGQQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNMY 60
DB 1 MMASKDPTNMDGTSAGQQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNMY 60
QY 61 VQAPQGEFTISNNTPGDIPLDQLGPHLNPFLSHLAQMYNGVGNMKVYKVLGNAFTA 120
DB 61 VQAPQGEFTISNNTPGDIPLDQLGPHLNPFLSHLAQMYNGVGNMKVYKVLGNAFTA 120
QY 121 GKIIISCTPPGAAQNIISQAQTMFPHVIAVRVLEPIEVPLEDVNRVLFHNNND-NAPTM 179
DB 121 GKIIISCTPPGAAQNIISQAQTMFPHVIAVRVLEPIEVPLEDVNRVLFHNNND-NAPTM 179
QY 121 GKIIISCTPPGAAQNIISQAQTMFPHVIAVRVLEPIEVPLEDVNRVLFHNNND-NAPTM 180
DB 121 GKIIISCTPPGAAQNIISQAQTMFPHVIAVRVLEPIEVPLEDVNRVLFHNNND-NAPTM 180
QY 180 RLVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVVEQTKPFSVNLPL 239
DB 180 RLVCMLYTPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVVEQTKPFSVNLPL 239
QY 240 NTLNSRVPSLIKSMVSRDHQVQFONGRVTLTGQLOQTTPTSASOLCKIRGSVFHANG 299
DB 240 NTLNSRVPSLIKSMVSRDHQVQFONGRVTLTGQLOQTTPTSASOLCKIRGSVFHANG 299
QY 239 SLSLSNRAPLPISMSGISPDNQSQVQFNGRCTLGRLVGTTPVSLSHVAKIRGT----S 294
DB 239 SLSLSNRAPLPISMSGISPDNQSQVQFNGRCTLGRLVGTTPVSLSHVAKIRGT----S 294
QY 300 GGNGYNLTLDGSPYHAFSPAPIGFDPDLGECDDHMEASPTTFQNTGVDVQKINVKQESA 359
DB 300 GGNGYNLTLDGSPYHAFSPAPIGFDPDLGECDDHMEASPTTFQNTGVDVQKINVKQESA 359
QY 295 NGVINLTLDGTPFPFPGAPIGFDPDLGGCDWHIN--MTQFGHSSQTQVDVDTTPTD 351
DB 295 NGVINLTLDGTPFPFPGAPIGFDPDLGGCDWHIN--MTQFGHSSQTQVDVDTTPTD 351
QY 360 FAPHLGTQADGLSDSVSNNTMIKLGWSPVSDGHRGDVDPVTPRGVSTLTERRAQLAP 419
DB 360 FAPHLGTQADGLSDSVSNNTMIKLGWSPVSDGHRGDVDPVTPRGVSTLTERRAQLAP 419
QY 352 FVPHLGSIQANGIG----SGNYGVLSWISPPSHPSGSQVDLWKIPNYGSSITEATHLAP 407
DB 352 FVPHLGSIQANGIG----SGNYGVLSWISPPSHPSGSQVDLWKIPNYGSSITEATHLAP 407
QY 420 PIYPGFGGAIVFMSDFPIAHGTNGLSVPCTTPQEFVTHFVNEQAPTRGEAALLHYLDP 479
DB 420 PIYPGFGGAIVFMSDFPIAHGTNGLSVPCTTPQEFVTHFVNEQAPTRGEAALLHYLDP 479
QY 408 SVYPPGFGVIVFMSKMP---GPGAYNLFCLLPQEVISHLASEQAPTVEAALLHYVDP 464
DB 408 SVYPPGFGVIVFMSKMP---GPGAYNLFCLLPQEVISHLASEQAPTVEAALLHYVDP 464
QY 480 DTHRNLFGEKLYPEGFMTCVPNSGSGTPQPLNGVVFVSVWSRYQLKPVGTAGPA-C 538
DB 480 DTHRNLFGEKLYPEGFMTCVPNSGSGTPQPLNGVVFVSVWSRYQLKPVGTAGPA-C 538
QY 465 DTGRNLGEFKAYPDGFLTCVENGASSGPQQLPINGVFVSVWSRYQLKPVGTASSARG 524
DB 465 DTGRNLGEFKAYPDGFLTCVENGASSGPQQLPINGVFVSVWSRYQLKPVGTASSARG 524
QY 539 RLGIIR 544
DB 525 RLGIIR 530

RESULT 2
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; PRIOR FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4

Query Match      9.5%; Score 276; DB 12; Length 671;
Best Local Similarity 26.4%; Pred. No. 2.6e-17; Indels 102; Gaps 17;
Matches 102; Conservative 54; Mismatches 128;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNMYVQAPQGEFTI- 70
DB 127 DGSITA----PEQGTWVGVIAPESAQMSTAADWATGKSVDSWK-----EAFPSFH 173
QY 71 -----SPNPTGDIPLDQLGPHLNPFLSHLAQMYNGVGNMKVYKVLGNAFTAAGTII 125
DB 174 TSVNWTSETQKILFKQSLGFLNLPYLEHLAKLVANVSGSIEVRFSGSGVFGGLAA 233

; LENGTH: 547
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-4

Query Match      9.5%; Score 276; DB 12; Length 547;
Best Local Similarity 26.4%; Pred. No. 1.9e-17; Indels 102; Gaps 17;
Matches 102; Conservative 54; Mismatches 128;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNMYVQAPQGEFTI- 70
DB 3 DGSITA----PEQGTWVGVIAPESAQMSTAADWATGKSVDSWK-----EAFPSFH 49
QY 71 -----SPNPTGDIPLDQLGPHLNPFLSHLAQMYNGVGNMKVYKVLGNAFTAAGTII 125
DB 50 TSVNWTSETQKILFKQSLGFLNLPYLEHLAKLVANVSGSIEVRFSGSGVFGGLAA 109
QY 126 SCIPGFG-AAQNIISQAQTMFPHVIAVRVLEPIEVPLEDVNRVLFHNNNDNAPTMRLVCM 184
DB 110 IVVPGVDVQVQSTMLQ---YPHVLFDAQVPEVIFCLPDLKSTLYHLMSOTDITSLVIM 166
QY 185 LY-----TPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFLVPPNVVEQTKPFSVNLPLN 240
DB 167 VYNDLINPYANDANSGC---IVT--VETKPGDFKFKHLLKPPG-----SMLTHG 211
QY 241 TLSNSRVPSLIKSMVSRDHQVQFONGRVTLTGQLOQTTPTSASOLCKIRGSVFHANG 300
DB 212 SIFSDILPKTSSLWIGNYWSIDITDF-----VIRPFVFOANR 248
QY 301 GNQYN-----LTLDGSPYHA-----PESPA-PIGFPDL---GEC--- 331
DB 249 HFDNFOETAGMSPTRFRPISVITTEQNGAKLGIVGATDVIVGIPDGWPDITIPGELIPA 308
QY 332 -DWHM-----EASPTTFQNTGVDVQK 350
DB 309 GDYAITNGTNDITATGVDYDADIK 334

RESULT 3
US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2

; LENGTH: 671
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match      9.5%; Score 276; DB 12; Length 671;
Best Local Similarity 26.4%; Pred. No. 2.6e-17; Indels 102; Gaps 17;
Matches 102; Conservative 54; Mismatches 128;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNMYVQAPQGEFTI- 70
DB 127 DGSITA----PEQGTWVGVIAPESAQMSTAADWATGKSVDSWK-----EAFPSFH 173
QY 71 -----SPNPTGDIPLDQLGPHLNPFLSHLAQMYNGVGNMKVYKVLGNAFTAAGTII 125
DB 174 TSVNWTSETQKILFKQSLGFLNLPYLEHLAKLVANVSGSIEVRFSGSGVFGGLAA 233

```

QY 126 SCIPGF-AAQNISIAQATMPHHVIADVRVLEPIEVPLEDVNRVLFHNNDNAPTMRVLCM 184  
DB 234 IWVPGVDPVQSTMLQ---YPHVLFDAQVEPVIFCLPDLRSTLYHLMSDITDTSVLM 290  
QY 185 LY----TPLRSGSSGCTDPFVIAGRVLTCPSPDFSLFVLPVPPNVEQTKFSPVNLPLN 240  
DB 291 VYNDLINPYANDANSSGC---IVT---VETKPGDPFKFHLKPPG-----SMLTHG 335  
QY 241 TLSNRVPSLIKMMVSRDHGMQVQFNGRVLTLDQLOQTTPTSASQCKIRGVSFHHANG 300  
DB 336 SIPSLIIPKTSLSLWIGNYWSIDTF-----VIRPFVQANR 372  
QY 301 GNGYN-----LTLDGSPYHA-----FESPA-PIGPPDL---GEC--- 331  
DB 373 HFDFNQETAGWSTPRFRPITITVSEKGSKLGIQVATDVIVGIPDGWPDPTTIPGELIPA 432  
QY 332 -DWM-----EASPTQFNTGVDVTK 350  
DB 433 GDYAITNGTNDITATGYDTADIK 458

## RESULT 4

US-10-209-507-2  
; Sequence 2, Application US/10209507  
; Publication No. US20030109033A1  
; GENERAL INFORMATION:

; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V  
; FILE REFERENCE: 454313-3151.2  
; CURRENT APPLICATION NUMBER: US/10/209,507  
; PRIOR FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: 09/617,594  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 669  
; TYPE: PRT  
; ORGANISM: Feline calicivirus

US-10-209-507-2

Query Match 9.4%; Score 272.5; DB 14; Length 669;  
Best Local Similarity 25.9%; Pred. No. 5.6e-17;  
Matches 106; Conservative 54; Mismatches 134; Indels 115; Gaps 19;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPMIMNYVQAPQGEFTI- 70  
DB 126 DGDSSI--TTPEQGLVGGVIAEPNAQMSAVADVATGKSDSEW-----ESFFSFH 174  
QY 71 ----SPNNTPGDILFDLQGLPHLPFLSHLAQMYNGVGMKVKVLLAGNAFTAGKIII 125  
DB 175 TSVMWSTSETQKILFKQSLGFLNPLYLHLSKLVVANGSVDRFSISGSGVGGKLA 234  
QY 126 SCIPGF-AAQNISIAQATMPHHVIADVRVLEPIEVPLEDVNRVLFHNNDNAPTMRVLCM 184  
DB 235 IWVPGVDPVQSTMLQ---YPHVLFDAQVEPVIFSLPDLRSTLYHLMSDITDTSVLM 291  
QY 185 LY----TPLRSGSSGCTDPFVIAGRVLTCPSPDFSLFVLPVPPNVEQTKFSPVNLPLN 240  
DB 292 VYNDLINPYANDANSSGC---IVT---VETKPGDPFKFHLKPPG-----S 331  
QY 241 TLSNRVPS--LIKSMV---SRDHGMQVQFNGRVLTLDQLOQTTPTSASQCKIRGVS 295  
DB 332 MLTHGSIFSDILPKSSSLWIGNRWSDITDF-----VIRPFV 368  
QY 296 FHANGNGYN-----LTLDGSPYHA-----FESPA-PIGFPD--LGE 330

DB 369 FOANRHFDFNQETAGWSTPRFRPITITVSEKGSKLGTGVATDVIVGIPDGWPDPTTIGE 428  
QY 331 CDWHEMESPPTQFNTGVDVTKINVKQBSAFAPHLGTIQADGLSDVSVNT 379  
DB 429 -----ELTPAGDYSI-----TNGSGNDIATANAYDSADVITNT 461

## RESULT 5

US-10-209-507-4  
; Sequence 4, Application US/10209507  
; Publication No. US20030109033A1  
; GENERAL INFORMATION:

; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V  
; FILE REFERENCE: 454313-3151.2  
; CURRENT APPLICATION NUMBER: US/10/209,507  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR FILING DATE: 09/617,594  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 668  
; TYPE: PRT  
; ORGANISM: Feline calicivirus

US-10-209-507-4

Query Match 9.4%; Score 272; DB 14; Length 668;  
Best Local Similarity 26.7%; Pred. No. 6.3e-17;  
Matches 107; Conservative 46; Mismatches 134; Indels 114; Gaps 20;

QY 23 PEANT-----AEP-ISMPEVAGATAAATAGQVNMIDPMIMNYVQAPQGEFTI----- 70  
DB 133 PEQGLVGGVIAEPNAQMSAVADVATGKSDSE-----W-----EAFPSHTSVN 177  
QY 71 -SPNNTPGDILFDLQGLPHLPFLSHLAQMYNGVGMKVKVLLAGNAFTAGKIIISCI 129  
DB 178 WSTSETQKILFKQSLGFLNPLYLHLSKLVVANGSVIEVRSISGSGVGGKLA 237  
QY 130 PGF-AAQNISIAQATMPHHVIADVRVLEPIEVPLEDVNRVLFHNNDNAPTMRVLCMLV-- 186  
DB 238 PGIDPVQSTMLQ---YPHVLFDAQVEPVIFITPDLRNSLYHLMSDITDTSVLM 294  
QY 187 --TPLRSGSSGCTDPFVIAGRVLTCPSPDFSLFVLPVPPNVEQTKFSPVNLPLN 244  
DB 295 LINPYANDANSSGC---IVT---VETKPGDPFKFHLKPPG-----SMLTH 334  
QY 245 SRVPS-LI---KSMVSRDH--GMVQFNGRVLTLDQLOQTTPTSASQCKIRGVS 299  
DB 335 GSIPSDILPKSSSLWIGNRWSDITDF-----VIKPFVFOAN 371  
QY 300 GNGYN-----LTLDGSPYH---AFESPAPIGFDPGLGCDWHMEASP 339  
DB 372 RHDFNQETAGWSTPRFRPITITVSEKGSKLGIQVATDSIVP-GIPD-----GMPDPTTIP 426  
QY 340 TTQFNTGVDVTKINVKQBSAFAPHLGTIQADGLSDVSVNTN 380  
DB 427 EKLTPAGDYAITNGGNDITTA-----ADYDGASIIKNNTN 462

## RESULT 6

US-10-282-122A-67793  
; Sequence 67793, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu

Db 4343 A-----APYTAADTTTPAAVANLAVSANGATLTGDCGAGATVTVRAPDGTVL 4389  
QY 350 KQINVKQESAF-----PHLGTTOADGLSDVSVNTNMIAKLGHWVSVDGHRGDV 399  
Db 4390 GNATVAADGHFVSLSPAAITGESLSVVQDAQAQNVSPAQNTAP-GALAFATP-----4442  
QY 400 DPWVTPRYGSTLTTEAAQLAPPIYPPGFGEAIVFFMSDFPIAHGTNGL---SVPCTIPOEF 456  
Db 4443 DNLILAADGLSVGTAEGSTIK-----VYGPNGVLLGSSPVTNDGTF 4485  
QY 457 VTFHVNQAPTGEAALLHYLDPD 480  
Db 4486 TVNLGSAQA--NGEVLQVSATGPD 4507  
RESULT 7  
US-10-369-493-2294  
; Sequence 2294, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; PRIOR FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2294  
; LENGTH: 913  
; TYPE: PRT  
; ORGANISM: Schizosaccharomyces pombe  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(913)  
; OTHER INFORMATION: unsure at all Xaa locations  
; US-10-369-493-2294  
Query Match 4.2%; Score 122; DB 15; Length 913;  
Best Local Similarity 19.5%; Pred. No. 0.037;  
Matches 116; Conservative 90; Mismatches 226; Indels 164; Gaps 30;  
QY 2 MMASKDAPTNDGTSAGQLVPEANTAEPIISME-----PVAGAATAAA-----TAG 47  
Db 238 IMVSVTPVAVGSTTSTVDPDVVFXAGNVEVYTPEDIMVSVTPVAVGSTTSTVDPDVVFXAG 297  
QY 48 --QVNMIDPMTMNNYVQAPQGEFTI-SPNNTPGDILFDL-OLGPHLNPFLSHLAQMNGW 103  
Db 298 IVEVNGIDPETIVSVTPVAVGSTTMTAPDEVSKXAVIVIDKETDPEIT--VSVTPVAVGST 355  
QY 104 VGNMKVKVLLAGNAP---TAGKIIISICPPCAQNIISIAQATMFRHVI-----ADVRVL 155  
Db 356 TSTVTRVVPVLAGNVEVYTPEDTMSVSVTPVAVGSTTSTV-----PEVPLAGNVEVVKVL 409  
QY 156 BP---IEVPLEBVRN-----VLFHNDNAPTMRVLCMLYTLFELRASGSSGSDTDPFVI- 203  
Db 410 EPEITMSVSVTPVAVGSTTSTVPEVPLAGNVEVNALEPDTMVSIVTPVAVGSTTSTVPEVVP 469  
QY 204 -AGRYLTCPSPDFSLFLVPPNVEQTKPFVSPN-LPLNTLSNRVPSLIKSMVMVSRDHG 261  
Db 470 LAGNVEVYTPEDTMSVSVTPVAVGSTTSTV-----TVPEVPL--AGNVEVYTPETETMVS-----521  
QY 262 QMWQONGREVTLGDLQGTTPTSASOLCKIRGSVFHANGGNGYNLTDLGSPYHAFESPA 321  
Db 522 -----VTVPA--VGSSTSTVPEVVPVLAG-----NVEVYTPETP-----ETMV 555  
QY 322 PIGFDPLGDCDWHMEASPTTQNTGDVIKQINVKQESAFPHLGTIOADGLSDVSVNTNM 381

; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 67793  
; LENGTH: 6310  
; TYPE: PRT  
; ORGANISM: Pseudomonas putida  
; US-10-282-122A-67793  
Query Match 4.2%; Score 122.5; DB 12; Length 6310;  
Best Local Similarity 19.5%; Pred. No. 0.67;  
Matches 110; Conservative 70; Mismatches 221; Indels 163; Gaps 22;  
QY 8 APTNMDGTSAGQLVPEANTAEPIISMEPVAGATAATAAGQVNMIDPMTMNNYVQAPQGE 67  
Db 4016 AAGNLSNLSAATVLSGGGEGAGASTVRDASGAILATGTVNQSQFOITLPSAQVTGSLQ 4075  
QY 68 FTIS-----PNNTEGDILFDLQGLPHLNPFLSHLAQMNG-----WVG 105  
Db 4076 VTLSDAAGNVSGPASLATPDHTPPAAI-----SNPVLSDQGRQLSGSGEAGATVQVR 4127  
QY 106 NMKVKVLGNAPFTAGKIIISICPPGPAQNIISIAQATMFRHVIADVRVLEIEVPLEDV 165  
Db 4128 NAAGALLGTATVGTGDRFTVTFDTPQANGQVIGVQMDAASNTSPAINVTPDLTPAPL 4187  
QY 166 RNVLFNHNDNAPTMRVLCMLYTPLRASGSS-----SGTDPFVI-----203  
Db 4188 TNNVLNNG-----LTLTGLGEAGATVTVHGPDTIIGTGLVAANGSFTLTLNSAQ 4238  
QY 204 -----AGR-----VLTCSPSPDFSLFLVPPNVEQTKP--PSVFNPLNTLSNS 245  
Db 4239 LNAQLLSVTQDAGNNTSTAVATAPDFT-----PP-----TAPTALALSGTGLQLTGNA 4288  
QY 246 RVPSLIKSMVMVSRD-HQOMVQFQNG--RVTL-----DQLOGTTPTTSASOLCKIRGSVFH 297  
Db 4289 EAGSTTVTRDASGNVLGTAVAGNGTFTQVTLNSAQTNQILQVLTATDAA-----GNVSP 4342  
QY 298 ANGGNGYNLTDLGDSPHYHAFESPAPIGFPDLGEC-----DWHMEASPTTQNTGDVI 349



Wed Jun 2 09:13:23 2004

```

; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 48055
; LENGTH: 1765
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-48055

Query Match      4.0%; Score 114.5; DB 12; Length 1765;
Best Local Similarity 20.9%; Pred. No. 0.55;
Matches 102; Conservative 58; Mismatches 198; Indels 129; Gaps 22;

QY      2 WKASKDAPN-----MDGTSGAGQLVPEANTAEPISE-----PVAGATAAAT-----AG 47
DB      966 MIAPTEPTSSVTPMTGAPG-----VVSPVTTEADGPVLPAASVAVTPTVPAG 1017
QY      48 QVNMDIPWIMNNYQAPQGEFTISNNTPGDILFDLQGLPHLNPPLSHLAQWYNGWGNM 107
DB      1018 S-----GVEGVVHAPLGSVTVASGVP-----SPFVSMLTVPASAVPDS 1058
QY      108 KVKVLLAGNAFTAGK-----IIISCIPGFAAQNISIAQATMPFHVIADVRVLEPIEVPLEDV 165
DB      1059 DVPSV-----AFTVGNAGAVVSIVSSKAVLGTLTLPAS-----VAVTVRLCAPSPKPVGV 1110
QY      166 RNVLPHNDNAPTRMLVCMLYT-----PLRAGSSSGTDPFVIAGRVLTCPSPDFS 216
DB      1111 -NVQFPGSAAVVPNNVPSYTLTLPASAVPLNVGSAVSSVLPPEMIA----- 1157
QY      217 FLFLVPNVQKTPSPVNPPLNTLSNRVPSLIKSMVSRDHGQWQVQNGRVTLDGQ 276
DB      1158 -----PTTEPTSSVTPMTGAPG-----AVSPVTTEADGPV-----LPAA 1194
QY      277 LQGTTPTSASQCKIRGVSFHANGNGYNLTLDGSPYHAFESPAPI-GFPDLGCDWHM 335
DB      1195 SVAVTPTVPAGSGVEGVVHAPLGSVTVASGVPSFPFVSMLTVPASAVPD----- 1246
QY      336 EASPTTQNTGDIKQINVKQESAFAPHLGTIOADGLSDVSVNTNMIK-----LGWVSP 390
DB      1247 NDVPSVAVTVGIAGAVSVIVRSKAV-----LGTLTLPASV-VAVTVRLCAPSPSAVGVNQV 1302
QY      391 VSDGHRGDVDPWIPRYGSLTTEAQ-----LAPPIYPFGGEAIVFFMSDFPIAHGTN 444
DB      1303 FPDGSAVVVPSNVVPSYTLTLPASAVPLNVGVASSVLP-----EATA-----PTTEPAS 1353
QY      445 GLSVPECT 451
DB      1354 SVTVPM 1360

```

```

RESULT 10
US-10-156-761-9262
; Sequence 9262, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089

```

```

; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9262
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9262

Query Match      3.9%; Score 113; DB 14; Length 724;
Best Local Similarity 20.5%; Pred. No. 0.19;
Matches 130; Conservative 48; Mismatches 220; Indels 236; Gaps 28;

QY      14 GTSAGAGQLVPEANTAEPISEMEPVAGATAAATAAGVNMID-----PWIMNNYVQ 62
DB      160 GDKIAVEAVSTQTVVDVADFEQVAAAAT--KPAGAVSVTDKGADPSGGGDSQTQAFRAIA 217
QY      63 APOGEFTISPNNTPGDILFDLQ-----WVGNMKVKVLLAGNAFTAGKIIISCIP 274
DB      218 AAGGVVWTP---PGDYRLTSSLNGVQNVTLQAGSMHVSVHTSRFIDQSSSSGGVHLKD 130
QY      92 FL-----SHLAQWYNG-----WVGNMKVKVLLAGNAFTAGKIIISCIP 322
DB      275 FAVIGEVTERVDSNPDNFVNGSLGPGSSVSGMWLQHLKVGLMLTGN-----ND 322
QY      131 GFAAQNISIAQATMPFHVIAD-----VRVLEPIEVPLEDVNRNVLPHNDNAPTR 180
DB      323 NLVVENNRILDT-----ADGLNLNGNARGVRVR-----NNFLRNGDDALAM- 365
QY      181 LVCMLYTPLRAGSSSGTDPFVIAGRVLTCPSPDFSFLVPPNVQKTPSPVNPPLN 240
DB      366 --NSLYSPDNTSSFSNT-----ISQPNLANG 390
QY      241 TLNSRVPSLIKSMVSRDH--GQWQFONGRVTLD--GQLOQTTPTSASQCKIRGVSF 296
DB      391 IAYGGTDLAVKKNLISDNTVALGSGIAISNOKE-LDPPSPLAGTITVSGNTL--VRAGAM 447
QY      297 HANGNGYNLTLDG-----SPHAFESAPICFPDPLGCDWHMEAS 338
DB      448 NPNQHPMGALRVDSYDSEANVSITGTTTSPYSAFEFVSGGGH---GVATKNTVVD 504
QY      339 PTTQNTGDIKQINVKQESAFAPHLGTIOADGLSDVSVN-----TNMIK 385
DB      505 GATVNTGTWV---VOAESQGAAKFSNVQRTGAGIYNCYPYAGSGTFTLTDCGGNA 560
QY      386 GWSPVSDGHRGDVDPWIPRYGSLTTE-----AAQLAPPIYPG---FGEAIV 431
DB      561 GWSST-----WGDCSAMPQCGQNPDPDPTRNLAKGRPATATGSDVYTFPKAVDGDAST 615
QY      432 FMSDFPIAHGTNGLSVPCITPQBFVTHFVNEQAPTRGEAALLHYLDPDTHRLNGEPKLY 491
DB      616 YWES-----TNN-----APPOAWTVDLGSSQAVRR---LVKLKPATAWQARTQTL 659
QY      492 PEG-----FMTCV-----PNSSGTGPOTLP 511
DB      660 VQGSTGSAISTVTVASQGYRFPDPATGNTATVTL 693

```

```

RESULT 11
US-10-128-714-3298
; Sequence 3298, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; TITLE OF INVENTION: Methods of Use
; FILE REFERENCE: 10182-018-999

```



```
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 3298
; LENGTH: 1555
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-3298

Query Match          3.9%; Score 113; DB 14; Length 1555;
Best Local Similarity 19.2%; Pred. No. 0.63;
Matches 104; Conservative 61; Mismatches 177; Indels 200; Gaps 28;

Qy 2 MMASKDAFTNMDGTSGAGQLVPEANTAPISMVPVAGAAATAAGQVNMIDPMIMNYV 61
Db 525 LKASTVSPAPL-GTSSAPTAPQKNFFELPLPPRPKSRPASSG-----RYTPNAPV 576

Qy 62 QAPQGEFTISP-----NNTPGDILFDLQGLHPLNPFSLHQAQMYNGVGNMKVKVLLAGN 116
Db 577 SAPLSQSIIPPANQYSNVPG-----APQSN-----IG----- 604

Qy 117 AFTAGKIIICIPPGFAAQNISIAQATMPFHVIADRVLEPIEVLVDVNVLFHNDNA 176
Db 605 -----PPD-----PPLOQPERLDPS-----NLLAPNVSA 631

Qy 177 PTMLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTPFSPVN 236
Db 632 PAVPSTASRYSP-RPPGVQAGVKP-----PPSPRYS-----PAPPQSTN--AVAA 673

Qy 237 LPLNTLSNSRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLOGTTPTPSASOLCKIRGSVF 296
Db 674 APRNRYASQ--PASISG-----QGAALQFO-----PRTSSPLA-YHEKIH 710

Qy 297 HANGGNGYLTDLGSPYHAFESAPIGFDPDLGECDMHMEASPTTQFNTG----DVIKQI 352
Db 711 YEDQCSERPOLQST-----ASPPLNHS-----HPSEQPVSSSENKSGVDVLENV 758

Qy 353 -----NVKQESAFAPHLGTIQADGLSDVSVNTNMIAKLGWVSPVSDG----- 394
Db 759 PPLSTRQSPKPNYPAP-----SAYTNEFANR--VAPVSTGPPPIAGMTGVLN 803

Qy 395 HRGVDVPMVIPRYGSTLTEAAQALAPPYPPGFGGAIIVFFMSDFPIAGHTNGLSVPTPIQ 454
Db 804 SSTEESPVPPRRSQTSQSPQLSPRLSVPS-----LDPFORPASVHGSTSP----- 850

Qy 455 EFVTHFVNEQAP-----TRGEAALLHYLDPDTHRNLGFEK-----LYPGEF----MTC 498
Db 851 ---TRTVNPYAPAPVPTHNRAPSVQLEFIPTDQGLDLSLRWKGAPIFKFGFGGAVISC 907

Qy 499 VP 500
Db 908 FP 909
```

## RESULT 12

```
US-10-128-714-8298
; Sequence 8298, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
```

```
; APPLICANT: Zamudio, Carlos
; APPLICANT: Eroshkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 8298
; LENGTH: 1832
; TYPE: PRT
; ORGANISM: Aspergillus fumigatus
US-10-128-714-8298

Query Match          3.9%; Score 113; DB 14; Length 1832;
Best Local Similarity 19.2%; Pred. No. 0.82;
Matches 104; Conservative 61; Mismatches 177; Indels 200; Gaps 28;

Qy 2 MMASKDAFTNMDGTSGAGQLVPEANTAPISMVPVAGAAATAAGQVNMIDPMIMNYV 61
Db 525 LKASTVSPAPL-GTSSAPTAPQKNFFELPLPPRPKSRPASSG-----RYTPNAPV 576

Qy 62 QAPQGEFTISP-----NNTPGDILFDLQGLHPLNPFSLHQAQMYNGVGNMKVKVLLAGN 116
Db 577 SAPLSQSIIPPANQYSNVPG-----APQSN-----IG----- 604

Qy 117 AFTAGKIIICIPPGFAAQNISIAQATMPFHVIADRVLEPIEVLVDVNVLFHNDNA 176
Db 605 -----PPD-----PPLOQPERLDPS-----NLLAPNVSA 631

Qy 177 PTMLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTPFSPVN 236
Db 632 PAVPSTASRYSP-RPPGVQAGVKP-----PPSPRYS-----PAPPQSTN--AVAA 673

Qy 237 LPLNTLSNSRVPSLIKSMVSRDHQMVQFQNGRVTLDGQLOGTTPTPSASOLCKIRGSVF 296
Db 674 APRNRYASQ--PASISG-----QGAALQFO-----PRTSSPLA-YHEKIH 710

Qy 297 HANGGNGYLTDLGSPYHAFESAPIGFDPDLGECDMHMEASPTTQFNTG----DVIKQI 352
Db 711 YEDQCSERPOLQST-----ASPPLNHS-----HPSEQPVSSSENKSGVDVLENV 758

Qy 353 -----NVKQESAFAPHLGTIQADGLSDVSVNTNMIAKLGWVSPVSDG----- 394
Db 759 PPLSTRQSPKPNYPAP-----SAYTNEFANR--VAPVSTGPPPIAGMTGVLN 803

Qy 395 HRGVDVPMVIPRYGSTLTEAAQALAPPYPPGFGGAIIVFFMSDFPIAGHTNGLSVPTPIQ 454
Db 804 SSTEESPVPPRRSQTSQSPQLSPRLSVPS-----LDPFORPASVHGSTSP----- 850

Qy 455 EFVTHFVNEQAP-----TRGEAALLHYLDPDTHRNLGFEK-----LYPGEF----MTC 498
Db 851 ---TRTVNPYAPAPVPTHNRAPSVQLEFIPTDQGLDLSLRWKGAPIFKFGFGGAVISC 907

Qy 499 VP 500
Db 908 FP 909
```

## RESULT 13

```
US-10-093-463-82
```

; Sequence 82, Application US/10093463  
; Publication No. US20030208039A1  
; GENERAL INFORMATION:  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Shenoy, Suresh  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: Gusev, Vladimir  
; APPLICANT: Pochart, Pascal  
; APPLICANT: Zhong, Mei  
; APPLICANT: Raestelli, Luca  
; APPLICANT: Mezes, Peter  
; APPLICANT: Smithson, Glenda  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Casman, Stacie  
; APPLICANT: Boldog, Ferenc  
; APPLICANT: Li, Li  
; APPLICANT: Zerhusen, Bryan  
; APPLICANT: Tchernev, Velizar  
; APPLICANT: Gangolli, Esha  
; APPLICANT: Vernet, Corine  
; APPLICANT: Pena, Carol  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Spytek, Kimberly  
; APPLICANT: Gorman, Linda  
; APPLICANT: Spaderna, Steven  
; APPLICANT: Voss, Edward  
; APPLICANT: Malyankar, Uriel  
; APPLICANT: Anderson, David  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Miller, Charles  
; APPLICANT: Taupier, Raymond J. Jr.  
; TITLE OF INVENTION: No. US20030208039A1e1 Antibodies that Bind to Antigenic Polypeptide  
; TITLE OF INVENTION: Encoding The Antigens, and Methods of Use.  
; FILE REFERENCE: 21402-230A (Cura 590AT)  
; CURRENT APPLICATION NUMBER: US/10/093,463  
; CURRENT FILING DATE: 2002-06-24  
; PRIOR APPLICATION NUMBER: 60/283,675  
; PRIOR FILING DATE: 2001-04-14  
; PRIOR APPLICATION NUMBER: 60/338,092  
; PRIOR FILING DATE: 2001-12-03  
; PRIOR APPLICATION NUMBER: 60/274,101  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/325,681  
; PRIOR FILING DATE: 2001-09-27  
; PRIOR APPLICATION NUMBER: 60/304,354  
; PRIOR FILING DATE: 2001-07-10  
; PRIOR APPLICATION NUMBER: 60/279,995  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 60/294,899  
; PRIOR FILING DATE: 2001-05-31  
; PRIOR APPLICATION NUMBER: 60/287,424  
; PRIOR FILING DATE: 2001-04-30  
; PRIOR APPLICATION NUMBER: 60/299,027  
; PRIOR FILING DATE: 2001-06-18  
; PRIOR APPLICATION NUMBER: 60/309,198  
; PRIOR FILING DATE: 2001-07-31  
; PRIOR APPLICATION NUMBER: 60/281,194  
; PRIOR FILING DATE: 2001-04-04  
; PRIOR APPLICATION NUMBER: 60/274,194  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/274,849  
; PRIOR FILING DATE: 2001-03-09  
; PRIOR APPLICATION NUMBER: 60/330,380  
; PRIOR FILING DATE: 2001-10-18  
; PRIOR APPLICATION NUMBER: 60/275,235  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: 60/288,342  
; PRIOR FILING DATE: 2001-05-03  
; PRIOR APPLICATION NUMBER: 60/275,578

; PRIOR FILING DATE: 2001-03-13  
; NUMBER OF SEQ ID NOS: 370  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 82  
; LENGTH: 2914  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-093-463-82  
  
Query Match 3.9%; Score 112.5; DB 15; Length 2914;  
Best Local Similarity 21.3%; Pred. No. 1.9; Mismatches 234; Indels 183; Gaps 33;  
Matches 132; Conservative 71;  
  
QY 5 SKDAPNTMDGTSGAQLVPEANTAEPISEMPVAGAATA---AATAGQVNMIDPWIMNIV 61  
DB 1473 SQGLTPSDEEPQLSQESFRTTHERP-ALTAAPLTTALNPVTAATEPVPVSPGQTOTL 1531  
QY 62 QAPQGEFTTSPNNTPGDILFDLQGLPHLNPFLSHLAQMVGWGNMKVKVLLAGNAFTAG 121  
DB 1532 QQPL-ELTAS-----QLPAGPTESP-----ASKGVTA 1558  
QY 122 KIIISCIPGFAAQNISIAQATMEPHVIADVRVLEPIEVPLEDVR-NVLFHNDN----- 175  
DB 1559 LLAIPHTP---ESSSLPVALQITPFGVSGA-----METTRVTIVIFAGSPNITVSS 1606  
QY 176 ---APTMLRLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFLFLVPPNVQKTKP 231  
DB 1607 RSPPARPFLMTKAVT-VRGHGS-----LPVRT-TPQPFSLT-----ASPSSRPVASP 1652  
QY 232 FSVNPLPLNTLSNSRVPSLIKSMVSR-----DHGMVQFQNGRVTLDDQLQGTTPPTS- 284  
DB 1653 GAISRSPTSSGSHKAVLTTPAVTKVISRTGVPOPTQOASASSPSTPLTVAGTAAEQVPVSP 1712  
QY 285 -ASQLCKIRGSVFHANGGNGYNLTLDSPYHAFESPAPIGFPDGLGECMDHWEASPTTQF 343  
DB 1713 LATRSLEIVLSTKEGAGHSQPM---GSP---ASFPQHPPLP-----SAPPRPAQ 1755  
QY 344 NTGDVIKQINVOESAPAPHLGTIQADG-----LSDVSVNTNM-IAKL 385  
DB 1756 HTTMATRSALPPTTAAASLST-ATDGLAATFMSLESTRPSQLLSGLPDTLSLAKV 1814  
QY 386 GWSPV-SDGHRGDVDWVPIRYGSTL-----TEAA--QLAPP--IYPPGFGEA 429  
DB 1815 GTSAPVATPGPKASVITTPLOQATTLPAQTLSPVLPTTAAAMTQAHPPTHIAPPAAGTA 1874  
QY 430 IVFFMSD-----FPIAHGTNGLSVECTIPQEFVTHFVNEQAPTGEAALLH----- 475  
DB 1875 PGLLLGATLPTSGVLPAEGT--ASMVSVVPRK-----STTKVAILSKQVSLPT 1922  
QY 476 -----YLDPDTHNLGEPFLYPEGFMW-----CVPNS---SGTGPQTLPINGVFV 517  
DB 1923 SMYGAEGGPTELTPTATSHPLTPLVAEPGAQAQATLVPVTSVLSRVSAARTAPQDSMLV 1982  
QY 518 FVSWVSRYQLKPVGT-AGP 536  
DB 1983 LLPQLAEAH-----GTSAGP 1997  
  
RESULT 14  
US-10-282-122A-46817  
; Sequence 46817, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert

```

Query Match      3.9%; Score 112; DB 12; Length 3930;
Best Local Similarity 18.8%; Pred. No. 3.4;
Matches 135; Conservative 86; Mismatches 264; Indels 232; Gaps 34;

Qy 5 SKDAP-----TNMDGT-----SGAGQLVPEANTAEPISMEPVAGAATAA-----43
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 746 SAUSPTFVDNPDGTTFTPINTFLINGVLQNNADPNVGVLPSIPANGSLTIVSYQVTTSL 805

Qy 44 -----ATAGQVNM-----DPWMN-----NVQAPQGSEFTI-----70
   ::|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 806 PTQNPTINSSSTQVSFILNPGDPPTIETSLNTVSTQINLANVVIKQVLDLTIADVQPI 865

Qy 71 -----SPNPTPGD--ILFD-LQLGPHLNP---FLSHLAQMYNGVGNMKVKVLLAGNA 117
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:

```

## RESULT 15

123	IIISCIPGFAAQNTISIAQATMFPHVIADRVLI--EPIEVPLEDVRNVLFHNNDNAPTWR	180
670	NNVSTTPRRPSTPRAAVTQ--TASQNAADVEWALRDQTAESPVEDSE-----	714
181	LVCMLYTPLRAGSSSGTDPFVIAGRVLTCPSPDFSEFLVPNPVEQKTKPESVP----	N 236
715	-----EEDDDSDTGSVSLGH--TTPSDYNDVISPSPQTEQSTPSIRKAKLS	763
237	LPLNTLSNRVPSLIKSMVMVRDHQM-----VFQNGRVTLDGLOQTTTPTSASQ	287
764	SPMTTSTSQKPVLGK--RVATPHASARAQTVTSTPVQ---GRV--EKVSGTSPVPAT	816
288	LKIRGSVFHANGCNGYNLTELDGSYPHAFESPAPIGFPDLGCDWHMEASPTT----	QFN 344
817	LLQ-----DQ-----ASSKTSRNV	834
345	TGDIVIKQINVKQBSAFAPHLGTTIQADGLSDVSVNTMI-----AKLGWSPVSD--GHR	396
835	SGARTSSASARQPSASASVLSPTEDDVSPVTSPLSMLSSASPSAKSPAPPSPKGRGR	894
397	GDVDPWVPIPRYGSITLFEAQLAPYV-----PGFGEAIVFPMDDPPIAHGTNGLSVPCT	451
895	VGV--PSLKPTLGG---KAVVGRPPSPVSGSAPGRLSGTSRAASTTPTYPAVTVVPPSS	950
452	IPOEFVTHFVNEQAPT---RGEAALLHYLDPDTHRLGEFK-----LYPEGFMTC	498
951	TAKSSVSNAFPVAPSPSLKPGSAAL-----QRRRSTGTAAGVSPVKSTTGMKTAVFDLS	1005
499	VPNSSGTGPQ	508
1006	SPKSGTGPQ	1015

Search completed: June 1, 2004, 14:04:26  
Job time : 36.8398 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.5159 Seconds  
(without alignments)  
4985.230 Million cell updates/sec

Title: US-09-926-799-1  
Perfect score: 2896  
Sequence: 1 MMASKDAPTMDGTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_78.\*  
1: Pir1.\*  
2: Pir2.\*  
3: Pir3.\*  
4: Pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1974.5	68.2	546	2 B37491	major capsid prote
2	1946	67.2	530	2 B37471	capsid protein - N
3	1223	42.2	542	2 S60616	capsid protein - h
4	1151.5	39.8	539	2 S40111	capsid protein - h
5	316	10.9	2344	1 R6W9H	genome polyprotein
6	315	10.9	576	2 A53982	capsid protein - E
7	311.5	10.8	2344	2 S55399	genome polyprotein
8	307.5	10.6	2344	2 S64740	genome polyprotein
9	300.5	10.4	702	1 A48582	coat protein - San
10	277.5	9.6	668	2 JQ2354	capsid protein - f
11	276	9.5	671	1 VCMWF9	coat protein - fel
12	274	9.5	668	1 VCMWF9	coat protein - fel
13	269	9.3	668	1 C48562	coat protein - San
14	264	9.1	703	1 C48562	capsid protein - f
15	261.5	9.0	668	2 JQ2356	genome polyprotein
16	147	5.1	2206	2 S03822	genome polyprotein
17	146.5	5.1	2205	1 GNNY2W	genome polyprotein
18	145.5	5.0	2207	1 GNNY5P	genome polyprotein
19	145	5.0	2206	1 GNNY4P	genome polyprotein
20	143.5	5.0	2194	1 GNNY87	genome polyprotein
21	141.5	4.9	2207	2 S09553	genome polyprotein
22	139.5	4.8	2206	1 GNNY27	genome polyprotein
23	137.5	4.7	2209	1 GNNY2P	genome polyprotein
24	136.5	4.7	2209	1 GNNY2P	genome polyprotein
25	135.5	4.7	2207	1 GNNY1P	genome polyprotein
26	133.5	4.6	2179	1 GNNY44	genome polyprotein
27	128.5	4.4	613	2 T35828	acetolactate synth
28	125	4.3	3473	1 A46112	genome polyprotein
29	125	4.3	3473	2 S27927	polyprotein - rice

genome polyprotein  
genome polyprotein  
polyprotein - infe  
hypothetical prote  
genome polyprotein  
genome polyprotein  
transcription fact  
genome polyprotein  
genome polyprotein  
genome polyprotein  
serum response fac  
large repetitive p  
probable RTX fami  
hypothetical prote

ALIGNMENTS

RESULT 1

B37491  
major capsid protein [similarity] - Southampton virus  
N:Alternate names: orf2 protein  
C:Species: Southampton virus  
C:Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 18-Aug-2000  
C:Accession: B37491

R:Lambdeny P8389 GATU-5300-1995  
Sequence 259, 516-519, 1995

A:Title: Sequence and genome organization of a human small round-structured (Norwalk-like virus, serotype 1)  
A:Reference number: A37491; MUID:93142023; PMID:8380940

A:Accession: B37491

A:Status: not compared with conceptual translation

A:Molecule type: nucleic acid

A:Residues: 1-546 <LAW>

A:Cross-references: GB:107418; NID:g1236787; PIDN:AAA92984.1; PID:g995114

A:Note: sequence extracted from NCBI backbone (NCBIP:123458)

A:Note: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype

C:Superfamily: human calicivirus capsid protein

C:Keywords: glycoprotein

F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 68.2%; Score 1974.5; DB 2; Length 546;  
Best Local Similarity 66.6%; Pred. No. 3.5e-139;  
Matches 367; Conservative 75; Mismatches 96; Indels 13; Gaps 7;

Qy 1 MMASKDAPTMDGTSGAGQLVPEANTAEPISEMPVAGATAAATAGQVNNIDPWIMNY 60

Db 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLMEPVAGPTTAVATAGQVNNIDPWIMNF 60

Qy 61 VQAPQGFSTISPNNTPGDILFDLQGLPHLPFLSLAQMYNGWGNMKVLLAGNAFTA 120

Db 61 VQSPQGFSTISPNNTPGDILFDLQGLPHLPFLSLAQMYNGWGNMKVLLAGNAFTA 120

Qy 121 GKIIISIPGFAAQNISIAQATMPPHVIADRVRLIEVPLEDVRNVLHNNDAPTMR 180

Db 121 GKIIIVCCVPGFTSSSLTIAQATLFPHVIADRVRLIEVPLEDVRNVLHNNDAPTMR 180

Qy 181 LVCMLYTLRASGSSGCTDFVITAGRVLTCPSPDFSLFLVPPNVEQTKPFSVPLNPLN 240

Db 181 LVCMLYTLPLRTGGSGNSDFVAVAGRVLTAPSSDFSLFLVPPNVEQTKPFSVPLNPLQ 240

Qy 241 TLNSRVPSLTKSMVSRDHGMVQFQNGRVTLDDQLQGTTPTSASOLCKIRGSVFHANG 300

Db 241 TLNSRVPSLTKSMVSRDHGMVQFQNGRVTLDDQLQGTTPTSASOLCKIRGSVFHANG 300

Qy 301 GNGYNLTDLGSPYHAFESPAFIGFDLGECDHMEASPT-TQFNTGDVIKQINVKOB-S 358

Db 299 ARTLNLTVDGKPPMAFDSAPVGFDFGKCDHMRISKTPNNTGSGDPMRESVSQTNVQ 358

Qy 359 AFAPHLGTIQADGLSDVSNTNMIAKLGWVSFVSDGHRGDVDPWIRYRGSTLTEAAOLA 418

Db 359 AFAPHLGTIQADGLSDVSNTNMIAKLGWVSFVSDGHRGDVDPWIRYRGSTLTEAAOLA 418

```

359 GFVPHLSIQFDEVFNHTG-DYIGTIEWISQSTPPTDINLWEIPDYGSSLSQAANLA 417
419 PPIVPPGGEAIVFMSDFPPTAHGTNGLS----VPCITPOEFVTHFVNEQAPTRGEAALL 474
418 PPVPPGGEALVYFVSAPP---GNNKRSAPNDVPCLLPQBYIITHFVSEQAPTWGDAALL 474
475 HYLDPDTHRNLFGEFKLYPEGFMTCVPNSSGTGPTQLPFGVVFVSVWSRFYQKLPVGTA 534
475 HYVDPDTHRNLFGEFKLYPGGYLTCVFNVGAGPQPLNGVFLFVSVWSRFYQKLPVGTA 534
535 GDA-CRLGIRR 544
535 STARGRLGVR 545

RESULT 2
B37471
capsid protein - Norwalk virus
C:Species: Norwalk virus
C:Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000
C:Accession: B37471
R:Yang, X.; Wang, M.; Wang, K.; Estes, M.K.
Virology 195, 51-61, 1993
A:Title: Sequence and genomic organization of Norwalk virus.
A:Reference number: A37471; MUID:93303939; PMID:839187
A:Accession: B37471
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: genomic RNA
A:Residues: 1-530 <JIA>
A:Cross-references: GB:M87661; NID:gl061311; PIDN:AAB50466.1; PID:gl061313
A:Note: sequence extracted from NCBI backbone (NCBIP:134157)
C:Superfamily: human calicivirus capsid protein

Query Match 67.2%; Score 1946; DB 2; Length 530;
Best Local Similarity 66.8%; Pred. No. 4.4e-137;
Matches 365; Conservative 67; Mismatches 96; Indels 18; Gaps 7;

QY 1 MMASKADAPNMDGTSGAGOLVPEANTAEPISEMPVAGATAAATAAGQVNMIDPMIMNY 60
DB 1 MMASKDATSSVDGAGAGOLVPEVNASDPLMDPVAGSSTAVATAGQVNPIDPMIINF 60
61 VQAPQGEFTISPNNTPGDILFDLQGLPHNPFSLHQAQMYNGVGMKVKVLLAGNAFTA 120
61 VQAPQGEFTISPNNTPGDVLFDLSLGLPHNPFLLHLSQMYNGVGMVGNVRIMLAGNAFTA 120
121 GKIIICIPGFAAQNTISQAQTMFPHVIADRVLEPIEVPLEDVNVLPHND-NAPTM 179
121 GKIIICIPGFGSHNLIAQATLPHVIADRVLEPIEVPLEDVNVLPHNDNRNQQT 180
180 RLVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVLPVNPVEQTKPFSVNLPL 239
181 RLVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVLPVNPVEQTKPFSVNLPL 238
240 NTLNSRVPSLIKMMVSRDHGMVQFQNGRVTLTGQLOQTTPTSASOLCKIRGVSF 299
239 SLSNRRAPLPISSIGISPDNVQVQFQNGRCLDGLVGTTPVSLSHVAKIRGT---S 294
300 GGNGVNLTELDGSPYHAFESAPITGFDLCECHMEASPTQNTGDVTKQINVKQESA 359
295 NGTVNLTELDGTPPHPEGPAPITGFDLCECHWHIN---WTQCHSSQTYDVIDTPTD 351
360 FAPHLGTIQAQGLSDSVSNVNMIAKLGWSPVSDGHRGDVDPWVPIRYGSLTTEAQLAP 419
352 FVPHLGSIQANGIG----SGNYVGVLSWISPPSPSGSQVDLWKIPNYGSSITEATHLAP 407
420 PIYPGFGAEIVFMSDFPIAHGTNGLSVPCITIQEFVTHFVNEQAPTRGEAALLHYLDP 479
408 SVYPPGFGELVFFMSKMP---GPGAYNLPLCLLPQBYIISHLASQAPTVGGAALLHYVDP 464
480 DTHRNLFGEFKLYPEGFMTCVPNSSGTGPTQLPFGVVFVSVWSRFYQKLPVGTA-PA-C 538
465 DTGRNLGEFKAYPDGFLICVFNAGSSGQQLPFGVVFVSVWSRFYQKLPVGTAASSARG 524

```

```

QY 539 RLGIRR 544
DB 525 RLGLRR 530

RESULT 3
S60616
capsid protein - human calicivirus (strain Melksham)
C:Species: human calicivirus
A:Variety: strain Melksham
C:Date: 23-May-1997 #sequence_revision 23-May-1997 #text_change 28-Jul-2000
C:Accession: S60616
R:Green, S.M.; Lambdon, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
Virus Res. 37, 271-283, 1995
A:Title: Capsid diversity in small round-structured viruses: molecular characterization
A:Reference number: S60615; MUID:96136658; PMID:8533462
A:Accession: S60616
A:Molecule type: genomic RNA
A:Residues: 1-542 <GRE>
A:Cross-references: EMBL:X81879; NID:g976077; PIDN:CAA57462.1; PID:g976079
A:Experimental source: strain Melksham
A:Note: it is uncertain whether Met-1 or Met-3 is the initiator
C:Superfamily: human calicivirus capsid protein
C:Keywords: capsid protein; coat protein

Query Match 42.2%; Score 1223; DB 2; Length 542;
Best Local Similarity 47.2%; Pred. No. 3.2e-83;
Matches 264; Conservative 73; Mismatches 188; Indels 34; Gaps 13;

QY 1 MMASKADAPNMDGTSGAGOLVPEANTAEPISEMPVAGATAAATAAGQVNMIDPMIMNY 60
DB 1 MKVASNDAAAPSTDGAAG---LVPESNN-EVMALEPVAGAAAPVTGQTNIIDPWIRANF 56
61 VQAPQGEFTISPNNTPGDILFDLQGLPHNPFSLHQAQMYNGVGMKVKVLLAGNAFTA 120
57 VQAPNGEFTVSPRNAPGCVLLNLELGPENLPYLAHLARMYNGVAGMEVQVMLAGNAFTA 116
121 GKIIICIPGFAAQNTISQAQTMFPHVIADRVLEPIEVPLEDVNVLPH-NNDNAPTM 179
117 GKLVAFAVPPHFFVENLSPOKITMFPVHVIDRVLEPIEVPLEDVNLSFFHYNQKDDPKM 176
180 RLVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFVLPVNPVEQTKPFSVNLPL 239
177 RIVAMLYTPLRNSG--SGDDVFTVSCRVLTRSPDFDFTVLPVPTVESKTKPFTLPILT 234
240 NTLNSRVPSLIKMMVSRDHGMVQFQNGRVTLTGQLOQTTPTSASOLCKIRGVS---F 296
235 GELSNSRFPVPIDQMYTSPNEVISVQCNQGRCLDGLQGTTLQLOVSGICAFKGEVTAHL 294
237 HANGG-NGYNLTDLGSPYHAFES-PAPIGFPD-----LGECMDHMEASPTQNTG 346
295 HDNHLNNVTITNLNGSFFDPSEDIAPLGVDPDFQGRVFGVISQORDKQNAAGHSEPANRG 354
347 -DVIKQINVKQESAFAPHLGTIQAQGLSDSVSNVNMIAKLGWSPVSDGHRGDVDPWVIP 405
355 HDVAVPTVTAQ---YTPKLGQIQIGTWQTDLLTVNPQVKF---TPVGLNTEHEFNQWVP 408
406 RYGSGLTTEAQLAPPIYPGFGAEIVFMSDFPIAHGTNGLSVPCITIQEFVTHFVNEQA 465
409 RYAGALNLTNLAAPSAPVFPGERLLFFRSULPLKGGYGNPAIDCLLPQEWVQHYQEAA 468
466 PTRGEAALLHYLDPDTHRNLFGEFKLYPEGFMTCVPNSSGTGPTQLPFGVVFVSVWSRF 525
469 PSKSEVALVRYINPTDGRALFEAKLHRAGFTVTSNTS--APVVVPANGYFRFDSWVNQF 526
526 YOLKPVGTAGPACRLGIRR 544
527 YSLAPWGTGN-----GRRR 540

RESULT 4
S40111
capsid protein - human calicivirus (strain Bristol isolate B493)

```

C;Species: human calicivirus  
A;Variety: strain Bristol isolate B493  
C;Date: 25-Dec-1994 #sequence\_revision 27-Feb-1997 #text\_change 28-Jul-2000  
C;Accession: S40111  
R;Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
submitted to the EMBL Data Library, December 1993  
A;Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de  
A;Reference number: S40111  
A;Accession: S40111  
A;Molecule type: genomic RNA  
A;Residues: 1-539 <GR>  
A;Cross-references: EMBL:X76716; NID:9436410; PIDN:CAA54134.1; PID:9436411  
A;Experimental source: human enteric calicivirus strain Bristol isolate B493  
C;Superfamily: human calicivirus capsid protein  
C;Keywords: capsid protein; coat protein

Query Match 39.8%; Score 1151.5; DB 2; Length 539;  
Best Local Similarity 44.9%; Pred. No. 6.8e-78;  
Matches 254; Conservative 78; Mismatches 185; Indels 49; Gaps 16;

Qy 1 MMASKDAPTMMDGTSGAGQLVPEANTAEPISEMEPVAGAAATAAGQVNMIDPWIMNYY 60  
Db 1 MKMASNDANPS-DGS--AANLVPEVNN-EVMALEPVVGAATAAPVAGQNVIDPWIRNF 56

Qy 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVLLAGNAFTA 120  
Db 57 VQAPQGEFTVSPNAPGEILMSAPLGDLPNLYSLHLSRMVNGYAGGFEVQVILAGNAFTA 116

Qy 121 GKIIISCIPTCPGFAAQNTSIAQATWPHVIAADVRLIEVLEPVEDVRLVLEH-NDNAPT 179  
Db 117 GKVIAFAVPPNFTPEGUSPSQVTPPHVIVDVRQLEPVLLIPDVRNNFYHYNQANDSTL 176

Qy 180 RLVCWLYTPLRASGSSGSDTDFVIAGRVLTCPSPDFSLFLVPPNVEOKTKPFVSNPLPL 239  
Db 177 KLIAMLYTPLRA--NNAGDDVFTVSCRVLTPSPDFDFILVPTVESRTKPFVPLVTV 234

Qy 240 NTLNSRVPSLIKSMVSRDHQVQFONGRVLTDLQGLTTPTSASOLCKIRGSVFHAN 299  
Db 235 EEMNSRPPIPLEKLYTGPSSAFVQVQNGRCTTGDVLLGTQLSANNICNFRGDVTHIA 294

Qy 300 GGNGY--NLTELDGSPYHAFES-PAPIGFPDL-----GECMDHMEASPTTQNT 345  
Db 295 GSHDYTNLASQNSNDYDTEIEIPALGTDPFGVKIOGLLTQTTTRADGSTRAHKAT-VST 353

Qy 346 GDVIKQINVKQESAFAPHLGTIQ--ADGLSDVSVNTNMIAKLGWVSPVSDG--HRGDVDP 401  
Db 354 GSV-----HFTPKLGSVQVTTDNDPQAGQN--TKFTPGVVIQDGDHQNRPQQ 401

Qy 402 WVIPRYGTLTEAQAQLAPPIYPGFGAIVFMSDFPIAHGTNGLSVCTPIPBQFVTHFV 461  
Db 402 WLLPNYSGRTGHNVLAPAVAPTPEGQLLFFRSTMPGCSGYNNMLDCLLPQEWVLHFY 461

Qy 462 NEQAPTRGEALLHYLDPTHNIGEPKLYPEGFWTCVPNSSGTGPQ--TLPINGVVFV 519  
Db 462 QEAAPAQSDVALLRFVNDTGRVLFECKLHKSGVITV---AHTGPVLDVLPPNGYPRFD 517

Qy 520 SWSRFYOLKPVGTAGPACRIGRS 545  
Db 518 SWNQFYTLAPMGNG-----TGRRA 538

RESULT 5  
RRWRH  
Genome polyprotein - rabbit hemorrhagic disease virus  
N;Contains: RNA-directed RNA polymerase (EC 2.7.7.48)  
C;Species: rabbit hemorrhagic disease virus  
C;Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 29-May-1998  
C;Accession: A41039  
R;Meyers, G.; Wirblich, C.; Thiel, H.J.  
Virology 184, 664-676, 1991  
A;Title: Rabbit hemorrhagic disease virus--molecular cloning and nucleotide sequencing c  
A;Reference number: A41039; MUID:91361557; PMID:1840711  
A;Accession: A41039

A;Molecule type: genomic RNA  
A;Residues: 1-2344 <MEY>  
A;Cross-references: GB:M67473  
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C;Keywords: nucleotidyltransferase

Query Match 10.9%; Score 316; DB 1; Length 2344;  
Best Local Similarity 28.3%; Pred. No. 9.6e-15;  
Matches 132; Conservative 60; Mismatches 172; Indels 102; Gaps 21;

Qy 12 MDGTSGAGQLVPEANTAEPISEMEPVAGAAATAA-----MEPVAGAAATAA-----ATAG----- 47  
Db 1766 MEGKARAAPGEAAAGTATTASVPTTTDGDMPGVVATTSVITAENSASINATAGIGGPPQ 1825

Qy 48 QVNMIDPWIMNYYVQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNM 107  
Db 1826 QVDOQETWRTNFYY---NDVFTMSVADAPGSILYTVQHSPPNNFTAVLSQMYAGWAGGM 1882

Qy 108 KVKVLLAGNAFTAGKIIISCIPPGFAAQNTSIA---QATWPHVIAADVRLIEVPLED 164  
Db 1883 QFREIVAGSGVFGGLVRAVIPPQ-----IEIGFGLVVRQPPHVVIDARSLEPVITIMP 1937

Qy 165 VRNVLEH-NDNAPTMRMLVCMLYTPL--RASGSSGSDTDFVIAGRVLTCPSPDFSLFLV 221  
Db 1938 LRPNMYHTPDGGLVPLTVLSVYNNLNPFGSIS-----AIQVTVETRPSEDFEFVIR 1992

Qy 222 PPNVEOKTKPFVSNPLPLNTLSNRVPSLIKSMVSRDHQVQFQ-----NGRVT 272  
Db 1993 APS--SKTVDSISPAGLLT-----PVLGTGNDNRWNGQIVGLQPVPGFSTCNRHN 2044

Qy 273 LDGLOQTTPTSASOLCKIRGSVPHANGGYNUTEL-----DGSFYHAFESP-----APIG 324  
Db 2045 LNSGTWSSPRFGDIDHRRGSASY-SGSNATNVLQFYANAGS---AIDNPISQVAPDG 2100

Qy 325 FPDLEGCDHMEASPTTQNTGDVIKQINVKQESAFAPHLGTIQADGLSDVSVNTNMIAK 384  
Db 2101 FPDMSFVFPNGGIPAAAGWVGFGAIWNSN-----SGAPNVTTVQA-----YE 2142

Qy 385 LGWVSPVSDGHRGDVDPWVIPRYGSTLTEAAQ-LAPPIYPGFGSEA 429  
Db 2143 LGPAT---GAPGNLQP-----TTNTSGAQTAKSIYAVVTGTA 2177

RESULT 6  
A53982  
capsid protein - European brown hare syndrome virus  
C;Species: European brown hare syndrome virus  
C;Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 28-Jul-2000  
C;Accession: A53982  
R;Wirblich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Eskens, U.; Haas, B.; Thiel, H.  
J. Virol. 68, 5164-5173, 1994  
A;Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease  
A;Reference number: A53982; MUID:94309183; PMID:7518531  
A;Accession: A53982  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-576 <WIR>  
A;Cross-references: GB:U09199  
C;Superfamily: human calicivirus capsid protein

Query Match 10.9%; Score 315; DB 2; Length 576;  
Best Local Similarity 26.2%; Pred. No. 1.5e-15;  
Matches 126; Conservative 69; Mismatches 200; Indels 86; Gaps 20;

Qy 7 DAP-----TNMDGTSGAGQLVPEANTAEPISEMEPVAGAAATAAGQVNMIDP---WTMN 58  
Db 8 DAPGTATTASVFGTGTGMDPGVAVSDVTVADNVAASVATAGTGGPQQASPOESNRVN 67

Qy 59 NYVQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVLLAGNAF 118  
Db 68 FFY---NDVFTMSVTDAPGSLYSVQHSPPNNFTQVLSQMYAGWAGMQPFRFVAGSGI 124

Qy 119 TAGKIIISCIPPGFAAQ-NISIAQATWPHVIAADVRLIEVPLEDVRNVLFH-NDNNA 176

Db 125 FGRLVCAIIIPGIIQPGLEVRQ---FPHVVIDARSLEPVTITMPDLRPEMHTGDPG 181  
Qy 177 PTMLRLVCMLYTFLRASGSSGTDPF-----VIAGRLVTCPSPDFSFLVLPNNVEQTKP 231  
Db 182 LVPTLWVSNNL-----INPFGTTSIAQVTVETRPSEDFEVLIRAPS--SKTVD 231  
Qy 232 FSVNPLPLNTLSNRVPSLIKSMVMVRDGHQMVQFQ-----NGRVTLDGOLQOTTP 282  
Db 232 SVNFSWLLTT-----PVLTCAGSDNRWGAIPVGLQPVPGGFSTSNRHNWNGSTYCWSS 285  
Qy 283 TSASQLCKIRGSVFHANGGYNLTDLGSPYHAFESP-----APIGPDLCGDWHMEAS 338  
Db 286 PRFDDIDHPSGNVSYPTGSATNTIETWYANAGTATTNPISNAPDGPDPMDGAIFP-----S 341  
Qy 339 PTTQNTGDVVIKQINVKQESAPAPHLGTIQADGLSDVSNTMIKLGWSPVSDGHRGD 398  
Db 342 GTT-IPTGAWVGFGQWNASNGTPYGVVQA-----YELGF-----ANGAPSS 383  
Qy 399 VDPWVPIPRYGSTLTAAGL-APPYIPPGFGE-----AIVFFMSDPPHAGTNGLSVPCTIP 453  
Db 384 IRP-----VTTTTGAQLVAKSIYGVIAQNSAGIIFLSKGMVS--TPGVAATTYP 434  
Qy 454 Q 454  
Db 435 Q 435  
RESULT 7  
S55399  
genome polyprotein - rabbit hemorrhagic disease virus (isolate B589)  
C:Species: rabbit hemorrhagic disease virus  
A:Variety: isolate B589  
C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 18-Jun-1999  
C:Accession: S55399  
R;Rossi, C.  
submitted to the EMBL Data Library, May 1995  
A:Reference number: S55399  
A:Accession: S55399  
A:Molecule type: genomic RNA  
A:Residues: 1-2344 <ROS>  
A:Cross-references: EMBL:X87607; NID:g854640; PIDN:CAA60910.1; PID:g854641  
A:Experimental source: isolate B589  
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C:Keywords: polyprotein  
Query Match 10.8%; Score 311.5; DB 2; Length 2344;  
Best Local Similarity 27.4%; Pred. No. 2.1e-14;  
Matches 127; Conservative 60; Mismatches 172; Indels 105; Gaps 20;  
Qy 17 GAGQLVPEANTAFISNEPVAGAAAT-----AAATAG-----QV 49  
Db 1768 GKARTAPQGEAAGTATTASVPGTTDGLDPGVATTSSVTAENSASATAGIGGPPQV 1827  
Qy 50 NMIDPIMNNVQAPOQEFTISPNNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGMKV 109  
Db 1828 DQQTWRTNFY---NDVFTWSVADAPGSLITYVQHSPPQNNPFTAVLSQMYAGWAGMGP 1884  
Qy 110 KVLLAGNAFTAGKIIISCIPGFAAQNISTA---QATMPFHVADRVRLPIEVPLEDYR 166  
Db 1885 RFIVAGSGVFGRLVAIVPPG-----IEIGPGLVEVRQFPHVVIDAKSLPEVTTITPDLR 1939  
Qy 167 NVLFH-NNDNAPTMRVLCMLYTPL--RASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPP 223  
Db 1940 PNMVHTPGDGLVPTLVLSVYNNLINPFGSGTS-----AQVTVETRPSEDFEFVIRAP 1994  
Qy 224 NVEQTKPFPSPVNLPLNTLSNSRVPSLIKSMVMVRDGHQMVQFQ-----NGRVTLD 274  
Db 1995 S--SKTVDISIPAGLLTT-----PVLTVGVNDNRNGQIVGLQFVPGGFSTCNRHWNLN 2046  
Qy 275 GLOQGTTPTSASQLCKTRGVSFHANGGNYNLTEL---DGSPVHAFESP-----APIGFP 326  
Db 2047 GSTYGWSSPRFADIDHRRGSASYP-GSNATNLQFHWANAGS---AVDNPISQVAPDGP 2102

Qy 327 DLGECDWHEASPTTQENTGDVVIKQINVKQESAPAPHLGTIQADGLSDVSNTMIKLG 386  
Db 2103 DMSFVFNPGPIPAAGVGFAGIWNNSN-----SGAPNVTTVQA-----YELG 2144  
Qy 387 WVSFVSDGHRGDVDPWVPIPRYGSTLTAAGL-APPYIPPGFGEA 429  
Db 2145 FAT----GAPGNLQP-----TTNTSGAQTVAKSIYAVVTGTA 2177  
RESULT 8  
S64740  
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)  
N:Contains: VP60 protein  
C:Species: rabbit hemorrhagic disease virus  
A:Variety: isolate AST/89  
C:Date: 12-Jul-1996 #sequence\_revision 26-Jul-1996 #text\_change 20-Jun-2000  
C:Accession: S64740; S46944; S49018; S65012  
R;Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.  
submitted to the EMBL Data Library, May 1995  
A:Description: Genomic organization of rabbit hemorrhagic disease virus determined by di  
A:Reference number: S64740  
A:Accession: S64740  
A:Molecule type: genomic RNA  
A:Residues: 1-2344 <CAS>  
A:Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552  
A:Experimental source: isolate AST/89  
R;Boga, J.; Casais, R.; Marin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.  
submitted to the EMBL Data Library, July 1993  
A:Description: Molecular cloning, sequence and expression of the capsid protein gene fro  
A:Reference number: S46944  
A:Accession: S46944  
A:Molecule type: genomic RNA  
A:Residues: 1650-2344 <BOG>  
A:Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623  
A:Experimental source: isolate AST/89  
R;Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.  
Virus Res. 27, 219-228, 1993  
A:Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus suppo  
A:Reference number: S49018; MUID:93255896; PMID:8488721  
A:Accession: S49018  
A:Molecule type: genomic RNA  
A:Residues: 1650-1796 <PAR>  
A:Cross-references: EMBL:Z24757  
A:Experimental source: isolate AST/89  
A:Accession: S65012  
A:Molecule type: protein  
A:Residues: 1767-1779; 1875-1877, 'X', 1879-1881; 1936-1938, 'X', 1940-1941 <PAW>  
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase  
C:Keywords: blocked amino end; polyprotein  
Query Match 10.6%; Score 307.5; DB 2; Length 2344;  
Best Local Similarity 26.3%; Pred. No. 4.1e-14;  
Matches 149; Conservative 67; Mismatches 212; Indels 139; Gaps 25;  
Qy 12 MDG---TSGAQLVPEANTAF-----PISMEPVAGAATAA-----ATAG----- 47  
Db 1766 MEKARTAPQGEAAGTATTASVPGTTDGLDPGVATTSSVTAENSASATAGIGGPPQ 1825  
Qy 48 QVNMIDPIMNNVQAPOQEFTISPNNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGM 107  
Db 1826 QVDQETWRTNFY---NDVFTWSVADAPGSLITYVQHSPPQNNPFTAVLSQMYAGWAGM 1882  
Qy 108 KVKVLLAGNAFTAGKIIISCIPGFAAQNISTA---QATMPFHVADRVRLPIEVPLED 164  
Db 1883 QPRFIVAGSGVFGRLVAIVPPG-----IEIGPGLVEVRQFPHVVIDAKSLPEVTTITP 1937  
Qy 165 VNVLFH-NNDNAPTMRVLCMLYTPL--RASGSSSGTDPFVIAGRVLTCPSPDFSLFLV 221  
Db 1938 LPNMVHTPGDGLVPTLVLSVYNNLINPFGSGTS-----AQVTVETRPSEDFEFVIR 1992  
Qy 222 PPNVQKTKPFPSPVNLPLNTLSNSRVPSLIKSMVMVRDGHQMVQFQ-----NGRVT 272



```
Db 1993 APS--SKTVDISPAGLIT-----PVLTVGNDNRNGQIVGLQVPVGGFSTCRHWN 2044
Qy 273 LDGQLQGTTPTSASOLCKIRGSVPHANGNGYNLTTEL-----DGSYPHAFESP-----APIG 324
Db 2045 LNSTYGVSSPRFGDIGHRRGSASYP-GNNATNVLQFWYANAGS---AIDNPISQVAPDG 2100
Qy 325 FPDLGECDDHMEASPTTQFNTGDIVKQINVKQESAFAPHLGTIQADGLSDSVNTNMIAX 384
Db 2101 FPDMSFVPFNGGIPPAAGWVGFGAIWNSN-----SGAENVTVQA-----YE 2142
Qy 385 LGWVSPVSDGHRGVDVPIPRYGSTLTERRAQ-LAPPIYPGFGELI-----VFFWSDPPI 439
Db 2143 LGPAT-----GAPGNLQ-----TTTSGSQTVAKSIYAVVTGTAQNPAGLFWASGVI 2191
Qy 440 AHGTNGLSVPCCTIQPFVTHFVNEQAPTRGEEAALLHYLDPTHRLNGLBEFKLYPGEFTCV 499
Db 2192 STPSANAITYPQPDRIVT-----TPGTAAAPVGNKTPIMFASVV 2232
Qy 500 PNSS-----GTGPQILPI 512
Db 2233 RRTGDVNATAGSANGTQYGTGSQLPV 2259

RESULT 9
A:8562
coat protein - San Miguel sea lion virus (serotype 1)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C:Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: A48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel
eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: A48562
A:Molecule type: genomic RNA
A:Residues: 1-702 <NEI>
A:Cross-references: GB:M87481; NID:G334882; PIDN:AAA16217.1; PID:G334884
A:Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBI:P:113565)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 10.4%; Score 300.5; DB 1; Length 702;
Best Local Similarity 25.2%; Pred. No. 2.4e-14;
Matches 137; Conservative 66; Mismatches 209; Indels 131; Gaps 25;

Qy 13 DGTSGAGQLVPEANTAPISMEPVAGATAAATAGQVNMID-PW-INNNYVQAPQGBFTI 70
Db 154 DPGGADIVTBEQGTVQQQVPQAQSALTTLTAAASTGKTVDCEWTFPSYHTA-----VNW 209
Qy 71 SPNTPGDILFDLOLGHNLNPLSHLAQMYNGVGNMKVLLAGNAFTAGKIILICIPP 130
Db 210 STTEAQKILFSRALSPELNPLRHISLSYSTSGGIDVRETVSGVGKGLAALIVPP 269
Qy 131 GFAAQNTISIAQATMFPPIADVRVLEPIEVPLEDVRNVLPHNNDAFTMRLVCMLYTPL- 189
Db 270 GI--EPVESPTMLQYPVLPFARQTEPIPIPIRKTLHSMDDTDTTLVIMVYNELI 327
Qy 190 -----RASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVVEQTKPFSVPN--LPPLNT 241
Db 328 NPYEQSEPKSSCSIT-----VETRPSSDFTFSLKPPG--SLLKHGSIPLIPRNS 377
Qy 242 --LNSRVPSLISKMMVSRDHGMVQFQNGR-----VTLD----- 274
Db 378 RHMGNRNMWSTIDGFV-----QPRVFSNRHFDFTSTTGWSTPYIPIEVLTKLDLRG 432
Qy 275 GQLQGTTPTSASQL-----CKIRGSVPHANGNGYNLTTEL-----DGSYPHAFESPA 321
Db 433 GQYKVTDTESLVPGLPDGHPDPTTIPMTASNGNVDYTVAEIRITNNGTHFGKFTMG 492
Qy 322 PI-----GFPDLGECDDHMEASPTTQFNTGDIVKQINVKQESAFAP-----HLGTIOA 369
```

```
Db 493 NLTTKVGKSNLGET-----QOTSRTLFLFASVG-----NYKQONTINPTHTKITSNLVVYDA 543
Qy 370 DGLSDSVSVNT-----NMIKLGWV-----SPVSDGHRGVDVDPWVPIPRYGSTLTTEAAQLA--P 419
Db 544 NVUSAATAKTITTTWISTHSLGVLVDESVPV-----GSDSTKVVRIATLP 587
Qy 420 PIYPGPGGELIIVFMSDFPIAHGNTGLSVPCCTIQPFVTHFVNEQAPTRGEEAALLHY-LD 478
Db 588 EAFTHG-GNPFVFTNKIQIGH-----FDRAHTKCFNSQVLMTOKLAENHYTLP 636
Qy 479 PDT 481
Db 637 PDS 639

RESULT 10
JQ2354
capsid protein - feline calicivirus (strain NADC)
C:Species: feline calicivirus
C:Date: 28-Aug-1985 #sequence_revision 07-Oct-1994 #text_change 20-Sep-1999
C:Accession: JQ2354
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.
J. Gen. Virol. 74, 2519-2524, 1993
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable
A:Reference number: JQ2354; MUID:94065683; PMID:7504075
A:Accession: JQ2354
A:Molecule type: mRNA
A:Residues: 1-668 <SEA>
A:Cross-references: GB:I09718; NID:G305104; PIDN:AAA16485.1; PID:G305105
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein

Query Match 9.6%; Score 277.5; DB 2; Length 668;
Best Local Similarity 23.3%; Pred. No. 1.2e-12;
Matches 124; Conservative 75; Mismatches 178; Indels 155; Gaps 24;

Qy 34 EPVAGATAAATAGQVNMIDPWIMNNYVQAPQGBFTI-----SPNTPGDILFDLOLQGP 87
Db 145 EPSAQMSTAADMATGKSVDSSEW-----EAFPSFHTSVNMWSTSETQKILFKQSLGP 195
Qy 88 HLANPFLSHLAQMYNGVGNMKVLLAGNAFTAGKIILICIPPGF-AAQNTISIAQATMFP 146
Db 196 LLNPYLEHLKLYVAVSGSVSEVRSISGSGVFGKLAIVVPPGVDVPVQSTMLQ---Y 252
Qy 147 HVIADVRVLEPIEVPLEDVRNVLPHNNDAFTMRLVCMLY-----TPLRASGSSGTDPFV 202
Db 253 HVLFDAQVDPVIFSIIDRLSTLYHLMPTDTTSLVIMVYNDLINPYNANDSNSSGC---I 309
Qy 203 IAGRVLTCPSPDFSLFLVPP----- 223
Db 310 VT--VETKPGDPKPFHLKPPGSMLTGHSVPSDLIPKSSSLWIGNRYWSDITDFVVRPFV 367
Qy 224 -----NVQKTKPFSVPLN-PLN-TLSNRVPSLISKMMVSRDHGMVQFQNGRVTLTD 274
Db 368 FOANRHFNFQETAGWSAPRFPITITISKSGKL--GIGVATDY-----IVP 414
Qy 275 GQLQGTTPTSASQLCKIRGSVPHANGNGYNLTDELGSPYHAFE-SPAPIGPPDLGECDD- 332
Db 415 GIPDGWPDITIAEDLTTPAGD-YAITSNGNDIT--TGSEYDSTVEIKNTNFRGMVYICGS 471
Qy 333 -----WHMEASPTTQF-----NTGDVIKQINVKQESAFAPHLGTIOADGLSD 374
Db 472 LQRAWGDKKISNTAFITTAIKEGNKIRPSNTIDMTK-IADVQDT---HVG-----EE 519
Qy 375 VSVNTNMIKLGWVSPVSDGHRGVDV-----WVPIPRYGSTLTTEAAQLAPPIYPGFGELI 430
Db 520 VQTSDDALLALTYGIGEQAIGSDRDRVVRISVLPEVGAR-----GNHNP 564
Qy 431 VFFWSDPPIAHGNTGLSVPCCTIQPFVTHFVNEQAPTRGEEAALLHY-LDPT 481
Db 565 IFYKNSIKUGVIRSIDV-----FNSQILHTSRQLSLNHYLLSPDS 605
```

VCWVFC  
coat protein - feline calicivirus (strain Japanese F4)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
C:Accession: B40481  
R:Tohya, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamaza  
Virology 183, 810-814, 1991  
A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.  
A:Reference number: A40481; MUID:91306470; PMID:1853578  
A:Accession: B40481  
A:Molecule type: Genomic RNA  
A:Residues: 1-668 <TOH>  
A:Cross-references: GB:D90357; NID:g221264; PIDN:BAA14371.1; PID:g221266  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:177,301,304,399,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.5%; Score 274; DB 1; Length 668;  
Best Local Similarity 31.0%; Pred. No. 2.1e-12;  
Matches 74; Conservative 39; Mismatches 82; Indels 44; Gaps 9;

QY 23 PEANTA-EPISMEPVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI-----SPNNT 75  
DB 133 PEQGTAVGVIAEPISAQMSAAMASGKSVDSW-----EAFPSHTSVNWSST 183  
QY 76 PGDILFDLQGLHNPFLSHLAQMYNGWGNMKVYLLAGNAFTAGKIIISCIPPGF-AA 134  
DB 184 QGKILFKQSLGPIILNPLYEHLKLYVWSSGIEVRFSGSGVFGKLAALVVPQVDPV 243  
QY 135 QNISIQAQTMPPHVIADRVLEPIEVLPEVLEVRNVLPHNDNAPTMLVCMY----TPLR 190  
DB 244 QSTSMQLQ---YPHVLFDAQVEPVIFCLPDLRSTLYHMSDSTDTTSLVIMVNDLINPYA 300  
QY 191 ASSGSSCTDFVITAGVILCPSPDFSLFLVPPNVQKTKPFSPVPLNTLSNSRVPS 249  
DB 301 NDSNSSGC---IVT--VETKPGDPFKHLKPPG-----SVLTGHSIPS 339

RESULT 13  
VCWVFF  
coat protein - feline calicivirus (strain CFI/68 FIV)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: A40507; B40507; T09246  
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.  
J. Virol. 65, 5440-5447, 1991  
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calicivirus.  
A:Reference number: A40507; MUID:91374597; PMID:1716692  
A:Accession: A40507  
A:Molecule type: Genomic RNA  
A:Residues: 1-668 <NEI>  
A:Cross-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875  
A:Accession: B40507  
A:Molecule type: protein  
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE>  
R:Neill, J.D.  
submitted to the EMBL Data Library, April 1998  
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.  
A:Reference number: Z16626  
A:Accession: T09246  
A:Status: preliminary, translated from GB/EMBL/DBJ  
A:Molecule type: Genomic RNA  
A:Residues: 1-668 <NE>  
A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256  
A:Experimental source: strain CFI/68 FIV  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.3%; Score 269; DB 1; Length 668;  
Best Local Similarity 31.0%; Pred. No. 5e-12;

VCWVFP  
coat protein - feline calicivirus (strain F9)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 28-Jul-2000  
C:Accession: B43382; C45538; FQ0407; S23702  
R:Carter, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.  
Virology 190, 443-448, 1992  
A:Title: The complete nucleotide sequence of a feline calicivirus.  
A:Reference number: A43382; MUID:92410623; PMID:1529544  
A:Accession: B43382  
A:Molecule type: Genomic RNA  
A:Residues: 1-671 <CARL>  
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
R:Carter, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.  
Arch. Virol. 122, 223-235, 1992  
A:Title: Identification and sequence determination of the capsid protein gene of feline calicivirus.  
A:Reference number: A45538; MUID:92117861; PMID:1731695  
A:Accession: C45538  
A:Molecule type: Genomic RNA; protein  
A:Residues: 1-671 <CAR2>  
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
A:Experimental source: strain F9  
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)  
R:Quiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.  
J. Gen. Virol. 73, 2429-2433, 1992  
A:Title: The cloning, sequencing and expression of a major antigenic region from the feline calicivirus.  
A:Reference number: FQ0407; MUID:93019069; PMID:1402818  
A:Accession: FQ0407  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 217-266, 'ST', 269-336, 'V', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>  
A:Cross-references: PIDN:AB23553.1; PID:g257083  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted  
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.5%; Score 276; DB 1; Length 671;  
Best Local Similarity 26.4%; Pred. No. 1.5e-12;  
Matches 102; Conservative 54; Mismatches 128; Indels 102; Gaps 17;

QY 13 DGTSGAGQLVPEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI- 70  
DB 127 DGSITA---PEQGTAVGVIAEPISAQMSAAMATGKSVDSW-----EAFPSFH 173  
QY 71 -----SPNNTPGDILFDLQGLHNPFLSHLAQMYNGWGNMKVYLLAGNAFTAGKII 125  
DB 174 TSNVWSTSETQKILFKQSLGPIILNPLYEHLKLYVWSSGIEVRFSGSGVFGKLA 233  
QY 126 SCIPPGF-AAQNISIQAQTMPPHVIADRVLEPIEVLPEVLEVRNVLPHNDNAPTMLVCM 184  
DB 234 IVVPPGVDPVQSTSMQLQ---YPHVLFDAQVEPVIFCLPDLRSTLYHMSDSTDTTSLVIM 290  
QY 185 LY-----TPLRAGSSGCTDFVITAGVILCPSPDFSLFLVPPNVQKTKPFSPVPLN 240  
DB 291 VYNDLINPYANDANSSGC---IVT--VETKPGDPFKHLKPPG-----SMLTHG 335  
QY 241 TLSNSRVPSLTKSMVNRDHQVQFQNGRVTLTGDLQGLQGTTPTSASOLCKIRGSVFHANG 300  
DB 336 SIPDLPTKTSLLIGNRYWSDITDF-----VIRPFVQANR 372  
QY 301 GNGYN-----LTELDSGYHA-----FESPA-PIGFPPDL---GEC--- 331  
DB 373 HFDENQETAGWSTFRFRPISVTITEQNGAKLIGVATDIYVPGIDGWPDTTTPGELIPA 432  
QY 332 -DWEM-----EASPTTQFNQGVN 350  
DB 433 GDYAITNGTGNDDITATGYDTADIIK 458

RESULT 12

Matches 74; Conservative 38; Mismatches 83; Indels 44; Gaps 9;

Qy 23 PEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI-----SPNNT 75  
Db 133 PEQGTWVGVIAPNAQNAQSTADMATGKSVDEW-----EAFPSFHTSVNMSTSET 183

Qy 76 PGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAPTAKGIIISCIPPGF-AA 134  
Db 184 QGKILFKOSGLPPLNPLYTHLAKLYVAMSGVDVRFSGSGVFGGKLAALVPPGIDPV 243

Qy 135 QNISTAQATMPHVIADVRVLEPIEVPLEDVRNVLFNHNDNAPTMRVLCMLY---TPLR 190  
Db 244 QSTWMLQ---YPHVLFDARQVEPVIFSPDURSTLYHLMSDITDTSVLVIMYNDLINPYA 300

Qy 191 ASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTKPFSVPLNLTLSNSRVPS 249  
Db 301 NDSNSSGC---IVT--VETKSGDPFKHLLKPPG-----SMLTHGSIPS 339

RESULT 14  
C48562  
coat protein - San Miguel sea lion virus (serotype 4)  
N:Alternate names: capsid protein  
C:Species: San Miguel sea lion virus  
C>Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
C:Accession: C48562  
R:Neill, J.D.  
Virus Res. 24, 211-222, 1992  
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lions.  
A:Reference number: A48562; MUID:92410750; PMID:1529644  
A:Accession: C48562  
A:Molecule type: Genomic RNA  
A:Residues: 1-703 <NEI>  
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888  
A:Note: sequence extracted from NCBI backbone (NCBIP:113567)  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.1%; Score 264; DB 1; Length 703;  
Best Local Similarity 27.4%; Pred. No. 1.3e-11;  
Matches 104; Conservative 53; Mismatches 181; Indels 42; Gaps 17;

Qy 13 DGTSGAGOLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTISP 72  
Db 154 DPGSAEIVTEQGTVVQQPAPAPTALATATATGKSVQEWTFP--SYHTSINWST 211

Qy 73 NNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAPTAKGIIISCIPPGF 132  
Db 212 VESQKILYSQLNPSINPYLDHIKLYSTWSGGIDVRFTVSGSGVFGGKLAALLVPPGV 271

Qy 133 -AAQNISTAQATMPHVIADVRVLEPIEVPLEDVRNVLFNHNDNAPTMRVLCMLYTP 191  
Db 272 EPIESVSMLO---YPHVLFDARQVEPVIFTPIDIRKTLFHSMDTDTTKLVINPY---E 324

Qy 192 SGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTKPFSVPLNLTLSNSRVPS 247  
Db 325 NGVENKTCSI---TVETRPADTFPALLKPPG--SLIKGSIPLDIPNSAHWMGNRW 379

Qy 248 PSLIKSMVMVRDQGMQVQNGR-VTLDGQLQG-TTPTASQLCKIRGVSFHANGGNG-- 303  
Db 380 WSTISGFSV-----QPRVFQSNRHFDFDSTTTGWSGTTPYVVFIEIKQKV---GSNNKW 430

Qy 304 YNLTELDGSPHAFESPAPIGFPDLGECDWMEASPTT-QNTGDKV--QINVKQESAF 360  
Db 431 FHVIDTD-----KALVPGIPDGWPTTIPD-----ETKATNGNFSYGESVRAGSTTIKPNENS 483

Qy 361 APLHGTIQADGLSDVSVNTN 380  
Db 484 THFKGTTCGLTSTVEIPEN 503

RESULT 15  
JQ2356  
capsid protein - feline calicivirus (strain KCD)  
C:Species: feline calicivirus  
C>Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
C:Accession: JQ2356  
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.  
J. Gen. Virol. 74, 2519-2524, 1993  
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable  
A:Reference number: JQ2354; MUID:94065683; PMID:7504075  
A:Accession: JQ2356  
A:Molecule type: mRNA  
A:Residues: 1-668 <SEA>  
A:Cross-references: GB:L09719; NID:g305107; PIDN:AAAL6487.1; PID:g305108  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein

Query Match 9.0%; Score 261.5; DB 2; Length 668;  
Best Local Similarity 31.1%; Pred. No. 1.8e-11;  
Matches 76; Conservative 37; Mismatches 86; Indels 45; Gaps 10;

Qy 19 GOLV-PEANT-AEPISEMPVAGATAAATAGQVNMIDPWIMNNYVQAPQGEFTI----- 70  
Db 128 GSVITPQGTWVGVIAPNAQNAQSTADMATGKSVDEW-----EAFPSFHTSVNM 178

Qy 71 SPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVKVLLAGNAPTAKGIIISCIPP 130  
Db 179 STSETQKILFKOSGLPPLNPLYTHLAKLYVAMSGSIEVRFSISGSGVFGGKLAALVVP 238

Qy 131 GF-AAQNISTAQATMPHVIADVRVLEPIEVPLEDVRNVLFNHNDNAPTMRVLCMLY--- 186  
Db 239 GVDPVQSTSMLO---YPHVLFDARQVEPVIFSPDLRSTLYHLMSDITDTSVLVMAYN 295

Qy 187 -TPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTKPFSVPLNLTLSNS 245  
Db 296 INPVANDSNSSGC---IVT--VETKSGDFFRHLKPPG-----SVLTHG 335

Qy 246 RVPS 249  
Db 336 SVPS 339

Search completed: June 1, 2004, 13:55:17  
Job time : 12.5159 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.58389 Seconds  
(without alignments)  
4310.252 Million cell updates/sec

Title: US-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPTNDGTSGAGQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1974.5	68.2	546	1	COAT_SOUV3
2	1148.5	39.7	539	1	COAT_LORDV
3	316	10.9	2344	1	POLN_RHDV
4	300.5	10.4	702	1	COAT_SMSV1
5	289	10.0	2208	1	POLN_MANCV
6	276	9.5	671	1	COAT_FCVF9
7	274	9.5	668	1	COAT_FCVF4
8	269	9.3	668	1	COAT_FCVF6
9	264	9.1	703	1	COAT_SMSV4
10	146.5	5.1	2205	1	POLG_POL2W
11	145.5	5.0	2207	1	POLG_POL2L
12	145	5.0	2206	1	POLG_POL3L
13	143.5	5.0	2194	1	POLG_HE701
14	139.5	4.8	2206	1	POLG_POL32
15	137.5	4.7	2209	1	POLG_POL1S
16	136.5	4.7	2208	1	POLH_POL1M
17	135.5	4.7	2206	1	POLG_POL1M
18	133.5	4.6	2179	1	POLG_HRV14
19	124.5	4.3	855	1	POLG_HRV3
20	124	4.3	2214	1	POLG_CXA24
21	123.5	4.3	2175	1	POLG_BOVEV
22	119.5	4.1	2196	1	POLG_EC05N
23	118.5	4.1	2194	1	POLG_EC30B
24	117.5	4.1	2185	1	POLG_CXB5P
25	117	4.0	788	1	BCSB_XANAC
26	117	4.0	2183	1	POLG_CXB4E
27	117	4.0	2183	1	POLG_CXB4J
28	116.5	4.0	2206	1	POLG_CXA21
29	115	4.0	2185	1	POLG_SVDVH
30	115	4.0	2185	1	POLG_SVDVU
31	113.5	3.9	2164	1	POLG_HRV69
32	112.5	3.9	2184	1	POLG_EC01F
33	112	3.9	2195	1	POLG_EC11G

## RESULT 1

COAT\_SOUV3  
ID COAT\_SOUV3 STANDARD; PRT; 546 AA.

AC Q04542;

DT 01-OCT-1994 (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE Coat protein (Capsid protein).

OS Southampton virus (serotype 3).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC ~~Neorovirus~~

OK NCBI\_TaxID=37129;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=93142023; PubMed=8380940;

RA Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;

RT "Sequence and genome organization of a human small round-structured

RT (Norwalk-like) virus."

RL ~~Sequence=2597516-539(49993)~~

CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see [http://www.ebi.ac.uk/announcements/](http://www.ebi.ac.uk/announcements) or send an email to [license@sib.ch](mailto:license@sib.ch)).

CC -----

DR EMBL; L07418; AAA92984.1; --

DR PIR; B37491; B37491.

DR InterPro; IPR004005; Calici\_coat.

DR InterPro; IPR008975; Viral\_Cap\_coat.

DR Pfam; PF00915; Calici\_coat; 1.

KW Coat protein; Glycoprotein.

FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;

Query Match 68.2%; Score 1974.5; DB 1; Length 546;

Best Local Similarity 66.6%; Pred. No. 2e-141;

Matches 367; Conservative 75; Mismatches 96; Indels 13; Gaps 7;

QY 1 MMASKDAPTNDGTSGAGQLVPEANTAEPTSMPEVAGAAATAAGVNMDIPWIMNY 60

DB 1 MMASKDAPQSDAGSAGAGQLVPEVTADPLPMEPVAGPTTAVATAGVNMDIPWVNF 60

QY 61 VQAPQGEFTTSPNTPGDILFDLQGLPHLPPLSHLAQMNGWGNMKVYLLAGNAFTA 120

DB 61 VQSPQGEFTTSPNTPGDILFDLQGLPHLPPLSHLSQMNGWGNMKVRVILLAGNAPSA 120

QY 121 GKIIISICPGFAQNISIAQATMPHVIADRVRLIEVLEPVEDVRVLFHNNNDNPTMR 180

DB 121 GKIIICVCPFGTSSSITIAQATLPHVIADRVRLIEPVEPVEDVRVLYHTNDNQPTMR 180

QY 121 GKIIICVCPFGTSSSITIAQATLPHVIADRVRLIEPVEPVEDVRVLYHTNDNQPTMR 180

DB 121 GKIIICVCPFGTSSSITIAQATLPHVIADRVRLIEPVEPVEDVRVLYHTNDNQPTMR 180

[illegible]

```

RESULT 2
COAT_LORDV STANDARD; PRT; 539 AA.
ID COAT_LORDV AC
PS4635;
01-OCT-1996 (Rel. 34, Created)
01-OCT-1996 (Rel. 34, Last sequence update)
01-NOV-1997 (Rel. 35, Last annotation update)
DT DT
DE Coat protein (Capsid protein).
OS Lordsdale virus (Human enteric Calicivirus).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
NC Norovirus.
NCBI_TaxID=82658;
[1]
RN SEQUENCE FROM N.A.
RX MEDLINE=96005060; PubMed=7561776;
RA Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;
RT "Human enteric Caliciviridae: the complete genome sequence and
expression of virus-like particles from a genetic group II small
round structured virus";
RT RT
RL J. Gen. Virol. 76:2349-2355 (1995).
CC - !- SIMILARITY TO THE CAPSID PROTEIN OF OTHER CALICIVIRUSES.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation
at the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; X86557; CAA60255.1; -.
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_Cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 539 AA; 58775 MW; 5E5C63E7F2C5FD21 CRC64;
Query Match 39.7%; Score 1148.5; DB 1; Length 539;
Best Local Similarity 44.9%; Pred. No. 4.1e-79;

```

	Matches	254;	Conservative	77;	Mismatches	186;	Indels	49;	Gaps	16;
Qy	1	MMMASKADPTNMDGTSGAGQLVP	EAETISMEPVAGAAATAATAGOVN	WIDPWIMNVY	60					
Db	1	MKMASNDANPS--DGS--	AANLVPEVNN--EYMALEPVVGA	IAAPVAGQQVNDPWIRKNN	56					
Qy	61	VOAPCGEFTISPNNTPGDILLFD	LQGLPHLPFLSHLAQMYNGWGN	MKVVLVLAGNAFTA	120					
Db	57	VOAPCGEFTVFRNAPGEILMSA	PLGPDLLPFLSHLSRMVNGYAG	GEVQVILLAGNAFTA	116					
Qy	121	GKIITSCIPGFAAQNTISIAQAT	MFPHVIADVRLVLEPIEVLPE	DRVNVLFH--NNDNAPT	179					
Db	117	GKVFIAAVPNFPFTEGLSPSQ	VTMFPHIIVDRQLPEVLIPLD	VRNMFVHYNQANDSTL	176					
Qy	180	RLVCMLYTPLRASGSSSTGDF	VFVIAGRVLTICPSPDFFLV	PPNVQKTKPSPVNLPL	239					
Db	177	KLIALMYTLPLRA--NNAGD	VFTVSCRVLTRPSPDFIFL	VPPTVFSRTPKFTVPVLTV	234					
Qy	240	NTLNSRVPSLIKSMVSRDHG	QWQVQFONGRVTLTGLOG	LTGTTPTTASOLCKIRGS	VFHAN	299				
Db	235	EMSNRSRPIPLEKLYTGFS	SAFVQVQNGRGTGVLGTT	QLTSVAVNICNFRG	VDVTHIA	294				
Qy	300	GGNGY--NLTEL	DGSPFYHAFES--PAPIG	FPDL-----GEC	DHMEASPTTQNT	345				
Db	295	GSHDYTNMLASQWNSYD	PTETEEIAPLGT	PDFVGKIQLGLLTQT	TRADGSTRAHKAT--VST	353				
Qy	346	GDVIQIKVYNQBSA	FAPHLGTTIQ--ADGL	SDSVNTNMIAKLGW	SPVSDG--HRGDVDP	401				
Db	354	GSV-----HFTPK	LGSVQFTTDTNND	FQAGQN--TKFT	PPGVQODGHHQNE	401				
Qy	402	WIPRYGSTLTEAAQLAP	PIYPGGEAIVFMSDF	PIAHGTNGLSV	PCTTIPQEEVTHFV	461				
Db	402	WSLPNYSGRTGRNVH	LAPAVATFPGEQLL	FFRSTWPGCSGY	PNMNLCLLPQEWLHFY	461				
Qy	462	NEQAPTRGEAALILH	LDPDTHENLGEFKLY	PEGFMTCPVNS	SGTGQP--TLPING	VFVVFV	519			
Db	462	QEAAPQAQSDVALL	RFPVNPDTGRVL	FECKLHKSGY	ITV----AHTG	PDLVLPNGYFRPD	517			
Qy	520	SWYSRFYOLKPVGT	AGACRLGIRRS	545						
Db	518	SWYNQFYLAPMG	-----TGRRA	538						

```

RESULT 3
POLN_RHDV
ID _POLN_RHDV STANDARD; PRT; 2344 AA.
AC
P27410;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Non-structural polypeptide protein [contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like
DE protein); Coat protein].
DE OS Rabbit hemorrhagic disease virus (RHDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Lagovirus.
NCBI TaxID=11976;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91361557; PubMed=1840711;
RA Meyers G., Wirblich C., Thiel H.-J.;
RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide
RT sequencing of a calicivirus genome.";
RL Virology 184:664-676 (1991).
CC -!- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE
CC -!- FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC

```









```
PX MEDLINE=91374597; PubMed=1716692;
RA Neill J.D., Reardon I.M., Heinrichson R.L.;
RT "Nucleotide sequence and expression of the capsid protein gene of
RL feline calicivirus.";
RL J. Virol. 65:5440-5447 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Neill J.D.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
-----
CC EMBL; M32819; AAA42925.1; -.
DR FIRM; U13992; AAC13993.1; -.
DR FIRM; A40507; VCMWFF.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 668 AA; 73550 MW; 9E52312108D4D441 CRC64;
SQ
Query Match 9.3%; Score 269; DB 1; Length 668;
Best Local Similarity 31.0%; Pred. No. 1.3e-12;
Matches 74; Conservative 38; Mismatches 83; Indels 44; Gaps 9;
QY 23 PEANT-AEPTSMPEVAGAAATAAGVNMIDPWNNVYVQAPQGEFTI-----SPNNT 75
DB 133 PEQGTWGVGVIAPNAQMSAADATCKSVDSSEW-----EAFSPHTSVNNWSTST 183
QY 76 PGDILFDLQGLPHNLFSLHQAQVNGVNMKVKLLAGNAFTAGKIIISCIIPGF-AA 134
DB 184 QGKILFQSLGLPLNPLTHLAKLYVANGSVDRFSISGSGVFGKLAIVVPGIDPV 243
QY 135 QNISIAQATMFPVHVIADRVLEIEVPLEDVRNVLPHNNDNAPTMLVCMLY-----TPLR 190
DB 244 QSTSMQLQ---YPHVLFDAQVPEVIFSDPLRSTLYHLSMDTDTTSLVIMVYNDLINPYA 300
QY 191 ASSGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQKTFPFPVNLPLNTLSNSRVPS 249
DB 301 NDSNSSGC---IVT--VETKPGDPFKPHLLKPPG-----SMLTHGSIPS 339
RESULT 9
ID COAT_SMSV4 STANDARD; PRT; 703 AA.
AC P36285;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Coat protein (Capsid protein).
OS San Miguel sea lion virus (serotype 4) (SMSV 4).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=36407;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92410750; PubMed=1529644;
RA Neill J.D.;
RT "Nucleotide sequence of the capsid protein gene of two serotypes of
RT feline calicivirus."
RT J. Virol. 65:5440-5447 (1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Neill J.D.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
-----
CC EMBL; M37482; AAA16220.1; -.
DR FIRM; C48562; C48562.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).
FT SEQUENCE 703 AA; 77721 MW; C5DAD8223B261073 CRC64;
SQ
Query Match 9.1%; Score 264; DB 1; Length 703;
Best Local Similarity 27.4%; Pred. No. 3.3e-12;
Matches 104; Conservative 53; Mismatches 181; Indels 42; Gaps 17;
QY 13 DGTSGAGQLVPEANTAEPISEMEPVAGAAATAAGVNMIDPWNNVYVQAPQGEFTISP 72
DB 154 DGPSSAEIVTEEQGTVVQQQAPAPATLATLATATAGTCKSVGEQWMTFF--SYHTSINWST 211
QY 73 NNTPGDILFDLQGLPHNLFSLHQAQVNGVNMKVKLLAGNAFTAGKIIISCIIPGF 132
DB 212 VESQGLIYSQALNPSINPYLDHIAKLYSTWSGIDVRETSGVFGKLAALVPPGV 271
QY 133 -AAQNTISIAQATMFPVHVIADRVLEIEVPLEDVRNVLPHNNDNAPTMLVCMLYTPLA 191
DB 272 EPIESVSMQLQ---YPHVLFDAQVPEVIFSDPLRSTLYHLSMDTDTTSLVIMVYNDLINPY 324
QY 192 SGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQKTFPFPVNLPLNTLSNSRV 247
DB 325 NGVENKTTCSI---TVETRFSAADFTFALLKPPG--SLIKHSIPSDLIIPNSAHMGNRW 379
QY 248 PSLLIKSMVSRDHQVQVQNGR--VTLDGQLQG--TTPTSASQLCKIRGSVFHANGNG-- 303
DB 380 WSTISGFSV-----QPRVFSQNRHDFDFTTTCGMSPTYYVPIEIKIQGV-----GSNNKW 430
QY 304 YNLTELDGSPYHAFESPAPIGFDPLGECWHMEASPTT-QFNTGDVVK--QINVKQESAF 360
DB 431 FHVITD-----KALVPGIPDGWPTTIPD---ETKATNGNFSYGESYRAGSTTIKPNENS 483
QY 361 APLHGTIQADGLSDVSNTN 380
DB 484 THFKGTVCITLSTVEIPEN 503
RESULT 10
ID POLG_POL2W STANDARD; PRT; 2205 AA.
AC P23069;
DT 01-NOV-1991 (Rel. 20, Created)
DT 01-NOV-1991 (Rel. 20, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2
DE (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Picornain 2A
DE (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein
DE P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28)
DE (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Poliovirus type 2 (strain W-2).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
```

Enterovirus.  
 NCBI\_TaxID=12085;  
 [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=90155230; PubMed=2154539;  
 Pevear D.C., Oh C.K., Cunningham L.L., Calenoff M., Jubelt B.;  
 "Localization of genomic regions specific for the attenuated, mouse-  
 adapted poliovirus type 2 strain W-2.";  
 J. Gen. Virol. 71:43-52(1990).  
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
 O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln|-Gly bond in the  
 poliovirus polyprotein. In other picornavirus reactions Glu may be  
 substituted for Gln, and Ser or Thr for Gly.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 (RNA) (N).  
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 each of which is composed of one copy each of proteins VP1, VP2,  
 VP3, and VP4.  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D06625; BAA00516.1; ALT\_SEQ.  
 DR PIR; A34032; GNNY2W.  
 DR HSSP; P03299; 1POV.  
 DR MEROPS; C03.001; -.  
 DR MEROPS; C03.020; -.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR004004; Calici\_pol\_hel.  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR001199; Pept\_3C\_picorn.  
 DR InterPro; IPR000081; Peptidase\_C3.  
 DR InterPro; IPR003138; Pico\_P1A.  
 DR InterPro; IPR002527; Pico\_P2B.  
 DR InterPro; IPR001676; RNv.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_P3vir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00548; Cys-protease-3C; 1.  
 DR Pfam; PF02226; Pico\_P1A; 1.  
 DR Pfam; PF00947; Pico\_P2A; 1.  
 DR Pfam; PF01552; Pico\_P2B; 1.  
 DR Pfam; PF00073; rhv\_3.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PR00918; CALICIVIRUSN.  
 DR ProDom; PD001125; Cys\_protease\_3C; 1.  
 DR ProDom; PD001306; Pico\_P2A; 1.  
 DR ProDom; PD001274; Pico\_P2B; 1.  
 DR SMART; SM00382; AAA; 1.  
 KW Polypeptide; Coat protein; Core protein; Transferase;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 KW Lipoprotein.  
 FT CHAIN 2 69 COAT PROTEIN VP4.  
 FT CHAIN 70 340 COAT PROTEIN VP2.  
 FT CHAIN 341 578 COAT PROTEIN VP3.  
 FT CHAIN 579 879 COAT PROTEIN VP1.  
 FT CHAIN 1028 1028 PROTEASE 2A.  
 FT CHAIN 1029 1125 CORE PROTEIN 2B.  
 FT CHAIN 1126 1454 CORE PROTEIN 2C.  
 FT CHAIN 1455 1541 CORE PROTEIN 3A.  
 FT CHAIN 1542 1563 GENOME-LINKED PROTEIN VP6.  
 FT CHAIN 1564 1746 PICORNAIN 3C.

FT	CHAIN	1747	2205	RNA-DIRECTED RNA POLYMERASE 3D. N-myristoyl glycine (in host) (By similarity).
FT	LIPID	2	2	
FT	ACT_SITE	1710	1710	PROTEASE (POTENTIAL).
FT	ACT_SITE	1724	1724	PROTEASE (POTENTIAL).
SO	SEQUENCE	2205 AA; 245701 MW; 2422AB039E0254AD CRC64;		

Query Match 5.1%; Score 146.5; DB 1; Length 2205;  
 Best Local Similarity 19.2%; Pred. No. 0.011;  
 Matches 109; Conservative 71; Mismatches 188; Indels 199; Gaps 26;

Qy	22	VPEANTAEPTSMPEVAGAAATAATAGQNNMIDPMNNYVQAPQGEFTISN-----NTP 76
Db	365	IPEFDVTPPID---IPGEVRNMMELAEIDTMIPLNTLSQRKNTVDMVRVELNDAHSPTP 421
Qy	77	GDILFDLQLGPHLPFLSH-----LAQWYNGWGNMKVKVLLAGNAFTAGKIIISCIPPG 131
Db	422	---ILCLSLSPASDPRLAHTMLGELINLYTHWAGSLKFTFLFCGSMATGKLLVYAPPG 478
Qy	132	FAAQNISIAQATMPPHVIADVRLERIEVPLEDVRLVFNH--NDNAP-----TMR 180
Db	479	AKAPE-SRKEAMLGTHVIDIGLQSSCTMVVPMISNTTYRQTINDSFTEGGYISMFYQTR 537
Qy	181	LVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFL-----VDPN----- 224
Db	538	VVPLSTPRKMD-----ILGFVSAC--NDFSRLLRDTHHSQEVNPOGLDLIE 585
Qy	225	--VBOKTKPFSVPLPLNTLSNR-----VPS-----LIKSMVYSRDHGQ 262
Db	586	GVVEGVTRNALTPLTPVNNLPDTRSSGPAHSKETPALTAVETGATNPLVPSDTVQTRH-- 643
Qy	263	MQV-----FQNGRVTLGQLQGTTP--SASQLCKIRGSVF----- 296
Db	644	VIQKRTRSESTVESFFARGACVAIEVDNDAPTTRASKLFSVMKITVYKDTQLARKLEFF 703
Qy	297	-----HANGGNGVNLTELDGSPHAFESPAIFGPDLG-----ECWHEA 337
Db	704	TYSRFDMFTFVTSNTYTDANNGHALNQVQIMY-----IPGAPIPKRNDYWTQTS 757
Qy	338	SPTTQFNTGDIKQINVKQBSAFAPHLGTTQA-----DGLSDSVSVNTMIAGLWVSPVS 392
Db	758	NPSVFYTYGAPPARISV-----PYGVIANAVSHFYDGFVKV----- 793
Qy	393	DGHRGDVDPWVIPRYGSTLTAQAAPPYPPGGEAIVFFMSDPPIAHNGTGLSVPTCI 452
Db	794	-----PLAQQAESTGDSL-----YGAA--SLNDF-----GSLAV----- 820
Qy	453	QSFVTHFVNEQAPTRGEAALLHYLDP 479
Db	821	-----RVVNDHNPTKLTISKIRVYMKP 841

RESULT 11  
 POLG\_POL2L STANDARD; PRT; 2207 AA.  
 ID POLG\_POL2L  
 AC P05210;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2  
 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Picornain 2A  
 (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein  
 P3A; Genome-linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28)  
 (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].  
 OS Poliovirus type 2 (strain Lansing).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 OX NCBI\_TaxID=12084;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86115399; PubMed=3003384;  
 RA la Monica N., Meriam C., Racaniello V.R.;  
 RT "Mapping of sequences required for mouse neurovirulence of poliovirus

RT type 2 Lansing ";

RL J. Virol. 57:515-525(1986).

CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN

CC O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.

CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the

CC poliovirus polyprotein. In other picornavirus reactions Glu may be

CC substituted for Gln, and Ser or Thr for Gly.

CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

CC {RNA} (N).

CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,

CC each of which is composed of one copy each of proteins VP1, VP2,

CC VP3, and VP4.

CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements/>

CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC

CC -----

CC EMBL; M12197; AAA46912.1; -.

CC PIR; A29507; GNNY5P.

CC HSSP; P03299; IPOV.

CC MEROPS; C03.001; -.

CC MEROPS; C03.020; -.

CC InterPro; IPR003593; AAA ATPase.

CC InterPro; IPR004004; Calici\_pol\_hel.

CC InterPro; IPR009003; Cys\_Ser\_trypsin.

CC InterPro; IPR000199; Pept\_3C\_picorn.

CC InterPro; IPR000081; Peptidase\_C3.

CC InterPro; IPR003138; Pico\_P1A.

CC InterPro; IPR002527; Pico\_P2B.

CC InterPro; IPR001676; Rnv.

CC InterPro; IPR006005; RNA\_helicase.

CC InterPro; IPR007095; RNA\_pol\_DS\_PS.

CC InterPro; IPR001205; RNA\_pol\_P3D.

CC InterPro; IPR007094; RNA\_pol\_PSVir.

CC InterPro; IPR008975; Viral\_Cap\_coat.

CC Pfam; PF00548; Cys-protease-3C; 1.

CC Pfam; PF02226; Pico\_P1A; 1.

CC Pfam; PF00947; Pico\_P2A; 1.

CC Pfam; PF01552; Pico\_P2B; 1.

CC Pfam; PF00073; rhv; 3.

CC Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.

CC Pfam; PF00910; RNA\_helicase; 1.

CC PRINTS; PR00918; CALICIVIRUSNS.

CC ProDom; PD001125; Cys\_protease\_3C; 1.

CC ProDom; PD001306; Pico\_P2A; 1.

CC ProDom; PD001274; Pico\_P2B; 1.

CC SMART; SM00382; AAA; 1.

CC Polyprotein; Coat protein; Core protein; Transferase;

CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;

CC Lipoprotein.

CC CHAIN 2 69 COAT PROTEIN VP4.

CC FT CHAIN 70 340 COAT PROTEIN VP2.

CC FT CHAIN 341 578 COAT PROTEIN VP3.

CC FT CHAIN 579 879 COAT PROTEIN VP1.

CC FT CHAIN 880 1028 PROTEASE 2A.

CC FT CHAIN 1029 1125 CORE PROTEIN 2B.

CC FT CHAIN 1126 1454 CORE PROTEIN 2C.

CC FT CHAIN 1455 1541 CORE PROTEIN 3A.

CC FT CHAIN 1542 1563 GENOME-LINKED PROTEIN VP6.

CC FT CHAIN 1564 1746 PICORNAIN 3C.

CC FT CHAIN 1747 2207 RNA-DIRECTED RNA POLYMERASE 3D.

CC FT LIPID 2 2 N-myristoyl glycine (in host).

CC FT ACT SITE 1710 1710 PROTEASE (POTENTIAL).

CC FT ACT SITE 1724 1724 PROTEASE (POTENTIAL).

CC SQ SEQUENCE 2207 AA; 245829 MW; 2BLE2070B7D44F99 CRC64;

Query Match 5.0%; Score 145.5; DB 1; Length 2207;

Best Local Similarity 18.8%; Pred. No. 0.013;

Matches 106; Conservative 72; Mismatches 193; Indels 193; Gaps 24;

QY 22 VPEANTAPISMEPVAGAAATAAATAGQVWIDPMNMYVQAPOGEFTISPN-----NTP 76

DB 365 IPEFDVTPEID---IPGEVRNMBELADITMPLNLTNQRKNTWDMYRVELNDAAHSDTP 421

QY 77 GDILFDLQGLHLPPLSH-----LAQWYNGWGNMKVKVLLAGNAFTAGKIIISCIPPG 131

DB 422 ---ILCLSLASPASDPLAHTMLGILNLYTHWAGSLKFTFLFCGSMATGKLLVSAPP 478

QY 132 FAAQNIASQATMPEPHVIADVRVLEPIEVPLEDVRLVFNH--NDNAP-----TMR 180

DB 479 AEAPK-SRKEAMLGTHVWDIGLQSSCTMVVFWISNTTYRQTINDSFTEGGYISMFYQTR 537

QY 181 LVCMLYTPLRASGSSGTDFFVIAGRVLTCPSPDPSPLF-----LVP 222

DB 538 VVVPLSTPRKMD-----ILGFVSAC--NDFSVELLRDTHISQEMPPQGLDLIE 585

QY 223 PNVEQKTKPFSVNPPLNLTLSNR-----VPS-----LIKSMVSRDHGQ 262

DB 586 GVVEGVTNALTPLTANNLPDTQSSGPAHSKETPALTAVETGATNPLVPSDTVQTRH-- 643

QY 263 MVQ-----FQGRVTLQGLQGTPTS-ASOLCKIRGSVF----- 296

DB 644 VIQKTRSESTVESFFARGACVAIIEVNDNDATPKRASKLFSWKITYKDTVQLRRKLEFF 703

QY 297 -----HANGGNGYNLTDLGSPYHAFESPAPIGPDPDCEGDWHMEASPT 340

DB 704 TYSRFDMETFTVTSNYTANNHALNQYQIMYPPGAPIGKWNV---TWOTSSNPS 760

QY 341 TQNTGDIKQINVKQESAFAPHLGTIQA-----DGLSDVSNTNMIAGLWVSPVSDGH 395

DB 761 VFYTGAPPARISV-----PYVGIANAYSHFYDGFYAKV----- 793

QY 396 RGVDPWVIIPRYGSTLTEAQLAPPIYPFGGEAIVFFMSDFPIAHGTNGLSVPTIPOE 455

DB 794 -----PLAQASTEGDSL-----YGAA--SLNDF-----GSLAV----- 820

QY 456 FVTHFVNEQAPTREGAALLHYLDP 479

DB 821 ---RVVNDHNPTKLTISKIRVYMKP 841

RESULT 12

POLG\_POL3L STANDARD; PRT; 2206 AA.

ID POLG\_POL3L 83; Q84783; Q84785; Q84786; Q84787; Q84788; Q84789;

AC P03302; Q84783; Q84785; Q84786; Q84787; Q84788; Q84789;

AC Q84790; Q84791; Q84792; Q84793; Q84794;

DT 21-JUL-1986 (Rel. 01, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins

DE P2A TO P2C, P3A; Genome-linked protein VP6; Picornain 3C

DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D

DE (EC 2.7.7.48)].

OS Poliovirus type 3 (strains P3/Leon/37 and P3/Leon 12A[1]B).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Enterovirus.

OX NCBI\_TaxID=12088;

RN [1]\_TaxID=12088;

RP SEQUENCE FROM N.A.

RC STRAIN=P3/Leon/37;

RX MEDLINE=84170338; PubMed=6324200;

RA Stanway G., Hughes P.J., Mountford R.C., Reeve P., Minor P.D.,

RA Schild G.C., Almond J.W.;

RT "Comparison of the complete nucleotide sequences of the genomes of

RT the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin

RT vaccine derivative P3/Leon 12a1b.";

RL Proc. Natl. Acad. Sci. U.S.A. 81:1539-1543 (1984).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=P3/Leon 12A[1]B;

RX MEDLINE=83299239; PubMed=6310508;  
 RA Stanway G., Cann A.J., Hauptmann R., Hughes P.J., Clarke L.D.,  
 RA Mountford R.C., Minor P.D., Schild G.C., Almond J.W.;  
 RT "The nucleotide sequence of poliovirus type 3 Leon 12 alb: comparison  
 RT with poliovirus type 1";  
 RL Nucleic Acids Res. 11:5629-5643(1983).  
 RN (3)  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-878.  
 RN MEDLINE=95120467; PubMed=7820548;  
 RA Grant R.A., Hiremath C.N., Filman D.J., Syed R., Andries K.,  
 RA Hogle J.M.;  
 RT Structures of poliovirus complexes with anti-viral drugs:  
 RT implications for viral stability and drug design.";  
 RL Curr. Biol. 4:784-797(1994).  
 RN (4)  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1-878.  
 RA Hiremath C.N., Grant R.A., Filman D.J., Hogle J.M.;  
 RT "Binding of the antiviral drug wins1711 to the Sabin strain of type-3  
 RT poliovirus - structural comparison with drug-binding in rhinovirus-  
 RT 14.";  
 RL Acta Crystallogr. D 51:473-489(1995).  
 CC -|- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
 CC O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
 CC -|- CATALYTIC ACTIVITY: Selective cleavage of Gln|-Gly bond in the  
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
 CC substituted for Gln, and Ser or Thr for Gly.  
 CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA}(N).  
 CC -|- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -|- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -|- MISCELLANEOUS: THE SEQUENCE OF STRAIN SABIN VACCINE P3/LEON/37 IS  
 CC SHOWN.  
 CC -|- MISCELLANEOUS: THE STRAIN SABIN VACCINE P3/LEON/37 IS THE  
 CC PROGENITOR OF THE STRAIN SABIN VACCINE P3/LEON 12A[1]B.  
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; K01392; AAA46914.1; -;  
 CC EMBL; X00925; CAA25444.1; -;  
 CC PDB; 1P1V; 03-JUN-95.  
 CC PDB; 1PVC; 15-SEP-95.  
 CC PDB; 1VBA; 11-JUL-96.  
 CC PDB; 1VBB; 11-JUL-96.  
 CC PDB; 1VBC; 11-JUL-96.  
 CC PDB; 1VBE; 11-JUL-96.  
 CC MEROPS; C03.001; -;  
 CC InterPro; IPR004004; Calici pol hel.  
 CC InterPro; IPR009003; Cys Ser trypsin.  
 CC InterPro; IPR000199; Pept\_3C\_picorn.  
 CC InterPro; IPR000081; Peptidase C3.  
 CC InterPro; IPR000138; Pico p1A.  
 CC InterPro; IPR002527; Pico p2B.  
 CC InterPro; IPR001676; Rhv.  
 CC InterPro; IPR000505; RNA helicase.  
 CC InterPro; IPR007095; RNA pol\_DS.  
 CC InterPro; IPR001205; RNA pol\_P3D.  
 CC InterPro; IPR007094; RNA pol\_P5vir.  
 CC InterPro; IPR008975; Viral cap coat.  
 CC Pfam; PF00548; Cys-protease-3C; 1.  
 CC Pfam; PF02226; Pico\_p1A; 1.  
 CC Pfam; PF00947; Pico\_P2A; 1.  
 CC Pfam; PF01552; Pico\_P2B; 1.  
 CC Pfam; PF00073; rhv; 3.

DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PD00918; CALICIVIRUSNS.  
 DR ProDom; PD001125; Cys\_protease\_3C; 1.  
 DR ProDom; PD001306; Pico\_P2A; 1.  
 DR ProDom; PD001274; Pico\_P2B; 1.  
 KW Polyprotein; Coat protein; Core protein; Transferase;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 KW 3D-structure; Lipoprotein.  
 FT CHAIN 2 69  
 FT CHAIN 70 340  
 FT CHAIN 341 578  
 FT CHAIN 579 878  
 FT CHAIN 879 1027  
 FT CHAIN 1028 1124  
 FT CHAIN 1125 1453  
 FT CHAIN 1454 1540  
 FT CHAIN 1541 1562  
 FT CHAIN 1563 1745  
 FT CHAIN 1746 2206  
 FT LIPID 2 2  
 FT ACT SITE 1709 1709  
 FT ACT\_SITE 1723 1723  
 FT VARIANT 431 431  
 FT VARIANT 864 864  
 FT VARIANT 908 908  
 FT STRAND 4 8  
 FT STRAND 25 29  
 FT STRAND 36 38  
 FT TURN 50 50  
 FT HELIX 51 54  
 FT STRAND 57 57  
 FT TURN 63 64  
 FT TURN 80 81  
 FT STRAND 83 87  
 FT TURN 88 89  
 FT STRAND 90 94  
 FT STRAND 101 102  
 FT HELIX 103 105  
 FT TURN 113 115  
 FT STRAND 123 123  
 FT HELIX 126 128  
 FT TURN 129 129  
 FT STRAND 133 134  
 FT STRAND 138 141  
 FT TURN 142 143  
 FT STRAND 147 151  
 FT TURN 152 152  
 FT HELIX 153 155  
 FT TURN 156 157  
 FT HELIX 159 167  
 FT STRAND 168 180  
 FT TURN 185 186  
 FT STRAND 187 197  
 FT TURN 198 198  
 FT STRAND 203 203  
 FT TURN 208 209  
 FT HELIX 213 216  
 FT TURN 219 221  
 FT STRAND 223 224  
 FT STRAND 226 226  
 FT TURN 235 236  
 FT STRAND 242 242  
 FT STRAND 245 245  
 FT HELIX 246 248  
 FT TURN 249 252  
 FT HELIX 255 260  
 FT STRAND 263 267  
 FT TURN 268 270  
 FT STRAND 273 278  
 FT STRAND 287 287  
 FT TURN 289 291  
 FT STRAND 292 292

FT	STRAND	295	307
FT	TURN	308	309
FT	STRAND	310	310
FT	STRAND	314	331
FT	STRAND	337	337
FT	TURN	348	351
FT	TURN	355	356
FT	STRAND	363	363
FT	TURN	366	367
FT	STRAND	379	380
FT	STRAND	382	382
FT	TURN	383	383
FT	HELIX	384	387
FT	TURN	388	388
FT	STRAND	391	392
FT	STRAND	397	397
FT	TURN	399	403
FT	HELIX	405	408
FT	STRAND	410	413
FT	TURN	414	415
FT	STRAND	418	419
FT	STRAND	423	426
FT	TURN	429	431
FT	TURN	433	437
FT	HELIX	439	444
FT	TURN	445	446
FT	STRAND	447	451

Query Match      5.0%; Score 145; DB 1; Length 2206;  
Best Local Similarity    20.5%; Pred. No. 0.015;  
Matches                  98; Conservative    58; Mismatches    194; Indels    128; Gaps    20;

QY	15	TSGAGQLVPEANTAPPIEMEPVAGATAAATAGOV-NM-----IDPWIMNNVQAPOG--	66
DB	347	TPGSNQYTSDNHQSFCAL-PEFDVTPIPIDPGEVKNMELAEIDTMIPLNESTKRTM	405
QY	67	--EFTISPNNTPGDILFDLQLGPHLPFLSH-----LAQMNGWGVGNMKVLLAGNAF	118
DB	406	DMYRVTLSDSADLSQPILCLSLSPASDPRLSHTMLGVLNYYTHWAGSLKFTFLFCGSM	465
QY	119	TAGKIICIPGFAAQNISIAQTAFHVIADVLEPIEVPLEDVNVLFH--NNDNA	176
DB	466	ATGKLIVAYPEG-AQPPTSRKEAMLGTHVIDWLGLSQSCINVVWISNVTVRQTQDSF	524
QY	177	PMRLVCMLY-----TPLRASSGSSGTDPFVIAGRVLTCPSDFSFLEL-----V	221
DB	525	TGGGIYSMEFYQRIIVVPLSTPKSMS-----MLGFVSAC--NDFSRLLRDTTHISQSAL	576
QY	222	PNVE-----QTKPFPVP---NLP-----LATLSNSRVPSLIKMM	255
DB	577	PGIEDLISEVAQGALLTSLPKQQDSLPTKASGPAHSKEVPALTAVETGATNPAPSOT	636
QY	256	VSRDHQMWQ-----FONGRVTLDGLOQOTTPTSASQLCKIRGSVFH-----	297
DB	637	VQTRH--VVQRRSRSESTIESFARGACVAIIIEVDNEQFTTRAQKLFAMWRITYKDTVOL	694
QY	298	-----ANGNGNYNLTELDGSPYHAFESPAGIPGPDIGECDM	333
DB	695	RKLEFFTYSRFDMFEFTFVVVTAFTNANGHALNQVYQIMYPGPAPTAKSMDDY--TW	751
QY	334	HMEASTPTQNTGDVIKQINVKQESAFAPHLGTIOA-----DGLSDVSVNNTNMAIKLG	386
DB	752	QTSNPNFSIFYGAAPARISV-----FVGLANAYSHEYDFGFAKVPLKTANDQIG	802

RESULT 13

POLG\_HE701

ID POLG HE701      STANDARD;      PRT;    2194 AA.

AC P32537;

DT 01-OCT-1993 (Rel. 27, Created)

DR 01-NOV-1995 (Rel. 32, Last sequence update)

DE 10-OCT-2003 (Rel. 42, Last annotation update)

DE Genome polypeptide [Contains: Coat protein VP4 (P1A); Coat protein VP2



```

FT ACT SITE 1723 1723 PROTEASE (POTENTIAL).
SQ SEQUENCE 2206 AA; 245731 MW; F226AD85403C37BA CRC64;

Query Match 4.8%; Score 139.5; DB 1; Length 2206;
Best Local Similarity 20.8%; Pred. No. 0.038;
Matches 96; Conservative 64; Mismatches 182; Indels 119; Gaps 21;

QY 22 VPEANTAEPISEPVAGAAATAAGVQNMIDPMNNYVQAPQGEFTI---SPNNTPGD 78
DB 365 IPEDVTPPID---IPGEVKNVMEIAEDITMIPLENLENTKNTMDYVRVLSDSANLSGP 421
QY 79 ILFDLQGLPHLNPFLSH-----LAQMYNGVGNMKVKVLLAGNAFTAGKIIISCIPGFA 133
DB 422 ILC-LSLSPAADPRLSHTMLGEVLNYYTHWAGSLKFTFLFCGSMWATGKLLVYAPPG-A 479
QY 134 AQNTLSIAQATWPFHVIAVRVLEPIEVLVDVNRVLFH--NNDNAPTMRVLVCMLY----- 186
DB 480 QPPTSRKAMLGTHVINDLGLQSSCTMVVPHISNVTYRQTQDSFTEGGYISMFYQTRIV 539
QY 187 TPLRASGSSGTDPEFVIAGRVLTCPSPDFSLFL-----VPPNVE-----QKT 229
DB 540 VPL-----STPKAMDMLGFVSAC--NDFSRLRLDRTTHISQAAMPQGVDDDLITEVAQNA 591
QY 230 KPFVSP-----NLP-----LNTLS-----NSRVPS 249
DB 592 LALSLEPKQSNLPDTPKASGPAHSKEVPTLTAVETGATNPLVPSDTVQTRHVIQORSSES 651
QY 250 LKSM-----MVSRDHQ-----MVQ--FQNGRVT-----LDQGLQGTPTTSASQLCK 290
DB 652 TIESFARGACVAIIIEVDNEQATNVQKLFATWRITYKDTVQLRRKLEFFTVSRFDMFT 711
QY 291 IRGSVFHANGNGYNLTELDGPYHAFSPAPIGFDPDLGECWHNEASPTTQFNTGDIK 350
DB 712 FVVTANFTNSNGHALNOVYQIMPPGAPTPEKSDWY---TWQTSNPSIFTYTGAAPA 768
QY 351 QINVKQESAFAPHLGTIOA-----DGLSDVSVNTNNIAKLG 386
DB 769 RISV-----PVGLANAYSHFYDGFVAKVPLKSDANDQVG 802

RESULT 15
POLG_POLLS STANDARD; PRT; 2209 AA.
ID FOLG_POLLS Q84881; Q84882; Q84883; Q84884; Q84885; Q84886; Q84887;
AC Q84888; Q84889; Q84890;
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DE Genome polypeptide (Contains: Coat protein VP4 (P1A); Coat protein VP2
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Picornain 2A
DE (P1A); Genome-linked protein VP3 (P1C); Core protein P2B; Core protein
DE P2A; Genome-linked protein VP3 (P1C); Picornain 3C (EC 3.4.22.28)
DE (EC 3.4.22.29) (P2A); Core protein P2B; Picornain 3C (EC 3.4.22.28)
DE Poliovirus type 1 (strain Sabin).
OS Poliovirus type 1 (strain Sabin).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OC NCBI_TaxID=12082;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=83299876; PubMed=6310545;
RA Nomoto A., Omata T., Toyoda H., Kuge S., Horie H., Kataoka Y.,
RA Genba Y., Nakano Y., Imura N.;
RT "Complete nucleotide sequence of the attenuated poliovirus Sabin 1
RT strain genome.";
RL Proc. Natl. Acad. Sci. U.S.A. 79:5793-5797(1982).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC O/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
CC poliovirus polypeptide. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
```

```

each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
-!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
-!- MISCELLANEOUS: THIS VIRUS IS A LIVE VACCINE STRAIN DERIVED FROM
THE MAHONEY STRAIN BY SPONTANEOUS MUTATIONS DURING THE ATTENUATION
PROCESS.
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

EMBL; V01150; CAA24465.1; -.
MEROPS; C03.001; -.
MEROPS; C03.020; -.
InterPro; IPR004004; Calici_pol_hel.
InterPro; IPR009003; Cys_ser_trypsin.
InterPro; IPR000199; Pept_3C_picorn.
InterPro; IPR000081; Peptidase_C3.
InterPro; IPR003138; Pico_P1A.
InterPro; IPR002527; Pico_P2B.
InterPro; IPR001676; Rhv.
InterPro; IPR000605; RNA_helicase.
InterPro; IPR007095; RNA_pol_DS_PS.
InterPro; IPR01205; RNA_pol_P3D.
InterPro; IPR007094; RNA_pol_P5vir.
InterPro; IPR008975; Viral_cap_coat.
Pfam; PF00548; Cys-protease_3C; 1.
Pfam; PF02226; Pico_P1A; 1.
Pfam; PF00947; Pico_P2A; 1.
Pfam; PF01552; Pico_P2B; 1.
Pfam; PF00073; rhv_3.
Pfam; PF00680; RNA_dep_RNA_pol; 1.
Pfam; PF00910; RNA_helicase; 1.
PRINTS; PR00918; CALICIVIRUSN.
ProDom; PD001125; Cys_protease_3C; 1.
ProDom; PD001306; Pico_P2A; 1.
ProDom; PD001274; Pico_P2B; 1.
Polyprotein; Coat protein; Core protein; Transferase;
RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
Lipoprotein.
CHAIN 2 69 COAT PROTEIN VP4.
CHAIN 70 341 COAT PROTEIN VP2.
CHAIN 342 579 COAT PROTEIN VP3.
CHAIN 580 881 COAT PROTEIN VP1.
CHAIN 882 1030 CORE PROTEIN P2A.
CHAIN 1031 1127 CORE PROTEIN P2B.
CHAIN 1128 1456 CORE PROTEIN P2C.
CHAIN 1457 1543 CORE PROTEIN P3A.
CHAIN 1544 1565 GENOME-LINKED PROTEIN VP3.
CHAIN 1566 1747 PICORNAIN 3C.
CHAIN 1748 2209 RNA-DIRECTED RNA POLYMERASE P3D.
LIPID 2 1712 N-myristoyl glycine (in host).
ACT SITE 1712 1712 PROTEASE (POTENTIAL).
ACT SITE 1726 1726 PROTEASE (POTENTIAL).
SQ SEQUENCE 2209 AA; 246576 MW; 9EC1EF4D174A28A4 CRC64;

Query Match 4.7%; Score 137.5; DB 1; Length 2209;
Best Local Similarity 19.3%; Pred. No. 0.054;
Matches 108; Conservative 73; Mismatches 193; Indels 187; Gaps 25;

QY 22 VPEANTAEPISEPVAGAAATAAGVQNMIDPMNNYVQAPQGEFTI---GDI 79
DB 366 LPEDVTPPID---IPGEVKNVMEIAEDITMIPDLSAKKNTMWRVRLSDKPHDDP 422
QY 80 LPFDLQGLPHLNPFLSH-----LAQMYNGVGNMKVKVLLAGNAFTAGKIIISCIPGFAA 134
DB 423 ILCSLSPASDPRLSHTMLGEILNYYTHWAGSLKFTFLFCGSMWATGKLLVYAPPG-AD 481
```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6

Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 31.7307 Seconds  
(without alignments)  
5419.275 Million cell updates/sec

Title: US-09-926-799-1

Perfect score: 2896

Sequence: 1 MMASKDAPTNMDGTSGAQ.....YQLKPVGTAGPACRLGIRRS 545

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL 25:\*

1: sp\_archaea:\*

2: sp\_bacteria:\*

3: sp\_fungi:\*

4: sp\_human:\*

5: sp\_invertebrate:\*

6: sp\_mammal:\*

7: sp\_mhc:\*

8: sp\_organelle:\*

9: sp\_phase:\*

10: sp\_plant:\*

11: sp\_rodent:\*

12: sp\_virus:\*

13: sp\_vertebrate:\*

14: sp\_unclassified:\*

15: sp\_rvirus:\*

16: sp\_bacteriap:\*

17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2873	99.2	545	12 Q8VA02	Q8VA02 human calic
2	2846	98.3	545	12 Q918C8	Q918C8 norwalk-lik
3	2846	98.3	545	12 Q9W182	Q9W182 norwalk-lik
4	2840	98.1	545	12 Q918C2	Q918C2 norwalk-lik
5	2798	96.6	545	12 Q91H12	Q91H12 human calic
6	2744	94.8	545	12 Q91V45	Q91V45 human calic
7	2605	90.0	544	12 Q66418	Q66418 desert shie
8	2139	73.9	543	12 Q8BCA3	Q8BCA3 human calic
9	2124	73.3	539	12 Q91V48	Q91V48 human calic
10	2017.5	69.7	543	12 Q8JXJ0	Q8JXJ0 norwalk-lik
11	1979	68.3	544	12 Q9DU46	Q9DU46 chiba virus
12	1977.5	68.3	546	12 Q8JW44	Q8JW44 norwalk-lik
13	1975	68.2	544	12 Q91I85	Q91I85 human calic
14	1974	68.2	544	12 Q9QIE7	Q9QIE7 chiba virus
15	1972.5	68.1	540	12 Q9YS14	Q9YS14 norwalk vir
16	1968	68.0	544	12 Q918C5	Q918C5 norwalk-lik

17	1964.5	67.8	546	12	Q9IV47	Q9IV47 human calic
18	1956.5	67.6	543	12	Q9IV43	Q9IV43 human calic
19	1956.5	67.6	544	12	Q8VA27	Q8VA27 human calic
20	1955.5	67.5	540	12	Q8BCA5	Q8BCA5 human calic
21	1954	67.5	544	12	Q918D1	Q918D1 norwalk-lik
22	1952	67.4	544	12	Q918J2	Q918J2 human calic
23	1946.5	67.2	543	12	Q918B9	Q918B9 norwalk-lik
24	1946	67.2	530	12	Q83884	Q83884 norwalk vir
25	1942	67.1	544	12	Q91V36	Q91V36 human calic
26	1941	67.0	544	12	Q91V41	Q91V41 human calic
27	1938	66.9	530	12	Q9QT39	Q9QT39 norwalk-lik
28	1930.5	66.7	531	12	Q8V7J5	Q8V7J5 norwalk-lik
29	1908.5	65.9	517	12	Q9JH41	Q9JH41 norwalk-lik
30	1904	65.7	539	12	Q8JVV5	Q8JVV5 norwalk-lik
31	1900	65.6	530	12	Q83876	Q83876 norwalk vir
32	1886	65.1	541	12	Q91V42	Q91V42 human calic
33	1288	44.5	522	12	Q8B558	Q8B558 bovine ente
34	1285	44.4	522	12	Q8B4Z5	Q8B4Z5 norwalk-lik
35	1276	44.1	522	12	Q8B4Z2	Q8B4Z2 norwalk-lik
36	1274	44.0	522	12	Q8V629	Q8V629 norwalk-lik
37	1269	43.8	522	12	Q8B4Y9	Q8B4Y9 norwalk-lik
38	1268	43.8	522	12	Q8V628	Q8V628 norwalk-lik
39	1261	43.5	520	12	Q7TBL1	Q7TBL1 bovine ente
40	1260.5	43.5	519	12	Q9YQ22	Q9YQ22 bovine cali
41	1256	43.4	520	12	Q7TBK7	Q7TBK7 bovine ente
42	1228	42.4	540	12	Q917X1	Q917X1 norwalk-lik
43	1223	42.2	542	12	Q66293	Q66293 calicivirus
44	1222	42.2	542	12	Q913B6	Q913B6 human calic
45	1219	42.1	540	12	Q91V50	Q91V50 human calic

## ALIGNMENTS

RESULT 1

Q8VA02 ID Q8VA02 PRELIMINARY; PRT; 545 AA.

AC Q8VA02; 01-MAR-2002 (TremBLrel. 20, Created)

DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)

DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)

DE Capsid protein.

OS Human calicivirus NLV/Potsdam 196/2000/DE.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.

OX NCBI\_TaxID=150135;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=HU/NLV/Potsdam 196/2000/DE;

RA Kunze, U., Schmees, E.,

RT Molecular epidemiology of outbreaks of gastroenteritis associated

RT with Norwalk-like viruses in Germany. "

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF439287; AAL32455.1; -

DR InterPro; IPR004005; Calici.coat.

DR InterPro; IPR008975; Viral\_cap.coat.

DR Pfam; PF00915; Calici.coat.1

SQ SEQUENCE 545 AA; 58601 MW; 7D47D4D9F3521D89 CRC64;

Query Match 99.2%; Score 2873; DB 12; Length 545;

Best Local Similarity 99.1%; Pred. No. 2.4e-230;

Matches 540; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPTSMSPVAGATAAATAGOVNMDPIMNMY 60

Db 1 MMASKDAPTNMDGTSGAGQLVPEANTAEPTSMSPVAGATAAATAGOVNMDPIMNMY 60

Qy 61 VQAPQGEFTTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120

Db 61 VQAPQGEFTTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMKVKVLLAGNAFTA 120

Qy 121 GKIIISICPGFAQAQNISIAQAATPFIADRVLEPIEVPLEDVRNVLFHNNONAPTMR 180

Db 121 GKIIISICPGFAQAQNISIAQAATPFIADRVLEPIEVPLEDVRNVLFHNNONAPTMR 180

Db	121	GKLIITCIPPGFAAQNTISIAQATMFPHVIADRVRLBPIEVPLEDVNRVLFHNNDNTPMR	180
Qy	181	LVCMLYTLPRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVBQKTKPFSVFNPLN	240
Db	181	LVCMLYTLPRASGSSGTDPEVIAGRVLTCPSPDFSLFLVPPNVBQKTKPFSVFNPLN	240
Qy	241	TLNSNRVPSLIKMMVSRDHQGMVQFQNGRVTLTGQLOGTTPTTSASQCKIRGSVFHANG	300
Db	241	TLNSNRVPSLIKMMVSRDHQGMVQFQNGRVTLTGQLOGTTPTTSASQCKIRGSVFHANG	300
Qy	301	GGNGYNLTLDGSPYHAFESPAPIGFDPDLGECOWMEASPTTQNTGDNVQINVKQESAF	360
Db	301	GGNGYNLTLDGSPYHAFESPAPIGFDPDLGECOWMEASPTTQNTGDNVQINVKQESAF	360
Qy	361	APHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWIPRYGSLTTEAAQLAPP	420
Db	361	APHLGTIQADGLSDVSVNTNMIKLGWSPVSDGHRGDVDPWIPRYGSLTTEAAQLAPP	420
Qy	421	IYPPGFGEAIVFMSDFPIAHGTNGLSVPCITPOEFVTHFVNEQAPTRGEAALLHVLDD	480
Db	421	IYPPGFGEAIVFMSDFPIAHGTNGLSVPCITPOEFVTHFVNEQAPTRGEAALLHVLDD	480
Qy	481	THRNLFGEFKLYPEGFMTCVPNSSGTGPOTLPINGVFVSVWSRYQLKPVGTAGPACRL	540
Db	481	THRNLFGEFKLYPEGFMTCVPNSSGTGPOTLPINGVFVSVWSRYQLKPVGTAGPACRL	540
Qy	541	GIRRS 545	
Db	541	GIRRS 545	
RESULT 2			
Qs18C8		PRELIMINARY;	PRT; 545 AA.
ID	Q918C8		
AC	Q918C8;		
DT	01-DEC-2001 (TrEMBLrel. 19, Created)		
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Capaid protein.		
GN	ORF2.		
OS	Norwalk-like virus NLV/Honolulu/219/1992/US.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_TaxID=171835;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Hu/NLV/Honolulu/219/1992/US;		
RX	MEDLINE=97193806; PubMed=9041391;		
RA	Ando T., Monroe S.S., Noel J.S., Glass R.I.;		
RT	"A one-tube method of reverse transcription-PCR to efficiently amplify		
RT	a 3-kilobase region from the RNA polymerase gene to the poly(A) tail		
RT	of small round-structured viruses (Norwalk-like viruses).";		
RL	J. Clin. Microbiol. 35:570-577 (1997).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Hu/NLV/Honolulu/219/1992/US;		
RX	MEDLINE=2026071; PubMed=1080147;		
RA	Ando T., Noel J.S., Fankhauser R.I.;		
RT	"Genetic classification of 'Norwalk-like viruses.'";		
RL	J. Infect. Dis. 181:S336-S348 (2000).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Hu/NLV/Honolulu/219/1992/US;		
RA	Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;		
RL	submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: AF414403; AAL12962.1;		
DR	InterPro: IPR004005; Calici coat.		
DR	InterPro: IPR008975; Viral Cap coat.		
DR	Fram; PF00915; Calici coat; 1.		
Qs	SEQUENCE 545 AA; 58573 MW; 760A0DD2AB9E5AD7 CRC64;		
Query Match			
Best Local Similarity 98.5%; Score 2846; DB 12; Length 545;			
Pred. No. 4.3e-226;			

Query Match 98.3%; Score 2846; DB 12; Length 545;  
Best Local Similarity 98.3%; Pred. No. 4.3e-228;  
Matches 536; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

QY 1 MMASKDPTNMDTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNY 60  
DB 1 MMASKDPTNMDTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNY 60

QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHNPFSLHQAQYNGWGNMKVKVLLAGNAFTA 120  
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHNPFSLHQAQYNGWGNMKVKVLLAGNAFTA 120

QY 121 GKIIISCIIPGFAAQNISIAQATMPHVIADRVLEPIEVLPLEDRVNNIDPWIMNY 180  
DB 121 GKIIISCIIPGFAAQNISIAQATMPHVIADRVLEPIEVLPLEDRVNNIDPWIMNY 180

QY 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSLFVPPNVQKTKPFSVPNLPLN 240  
DB 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSLFVPPNVQKTKPFSVPNLPLN 240

QY 241 TLSNRVPSLIKSMVSRDHQVQFQNGRVTLTGQGLTPTTSASQLCKIRGSVFHANG 300  
DB 241 TLSNRVPSLIKSMVSRDHQVQFQNGRVTLTGQGLTPTTSASQLCKIRGSVFHANG 300

QY 301 GNGYNLTDLGSPYHAFESPAPIGFDLGECDHMEASPTTQFNTGVDVIKQINVKQESAF 360  
DB 301 GNGYNLTDLGSPYHAFESPAPIGFDLGECDHMEASPTTQFNTGVDVIKQINVKQESAF 360

QY 361 APHLGTIOADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVPIPRYGSTITEAAQLAPP 420  
DB 361 APHLGTIOADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVPIPRYGSTITEAAQLAPP 420

QY 421 IYPGFGAIVFFMSDRPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480  
DB 421 IYPGFGAIVFFMSDRPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480

QY 481 THNLGFEKLYPEGFMTCPVNSSGTGPTLPINGVVFVSVSRFYQLKPVGTAGPACRL 540  
DB 481 THNLGFEKLYPEGFMTCPVNSSGTGPTLPINGVVFVSVSRFYQLKPVGTAGPACRL 540

## RESULT 4

QY18C2 PRELIMINARY; PRT; 545 AA.  
AC QY18C2;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Capsid protein.  
GN ORF2.  
OS Norway-like virus NLV/Little Rock/316/1994/US.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=171837;

[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Little Rock/316/1994/US;  
RX MEDLINE=97193806; PubMed=9041391;  
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
of small round-structured viruses (Norwalk-like viruses).";  
RL J. Clin. Microbiol. 35:570-577(1997).  
RN [2]

RP SEQUENCE OF 100-192 FROM N.A.  
RC STRAIN=Hu/NLV/Little Rock/316/1994/US;  
RX MEDLINE=96071277; PubMed=9407386;  
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,

RA Seto Y., Monroe S.S., Glass R.I.;  
RT "Correlation of patient immune responses with genetically  
characterized small round-structured viruses involved in outbreaks of  
nonbacterial acute gastroenteritis in the United States, 1990 to  
1995.";  
RL J. Med. Virol. 53:372-383(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Little Rock/316/1994/US;  
RX MEDLINE=20266071; PubMed=10804147;  
RA Ando T., Noel J.S., Fankhauser R.L.;  
RT "Genetic classification of 'Norwalk-like viruses.'";  
RL J. Infect. Dis. 181:S336-S348(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Little Rock/316/1994/US;  
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF414405; AAL12968.1; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 545 AA; 58666 MW; F55076D61D12E82 CRC64;

Query Match 98.1%; Score 2840; DB 12; Length 545;  
Best Local Similarity 98.5%; Pred. No. 1.4e-227;  
Matches 537; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 MMASKDPTNMDTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNY 60  
DB 1 MMASKDPTNMDTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNY 60

QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHNPFSLHQAQYNGWGNMKVKVLLAGNAFTA 120  
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHNPFSLHQAQYNGWGNMKVKVLLAGNAFTA 120

QY 121 GKIIISCIIPGFAAQNISIAQATMPHVIADRVLEPIEVLPLEDRVNNIDPWIMNY 180  
DB 121 GKIIISCIIPGFAAQNISIAQATMPHVIADRVLEPIEVLPLEDRVNNIDPWIMNY 180

QY 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSLFVPPNVQKTKPFSVPNLPLN 240  
DB 181 LVCMLYTPLRASGSSGTDPPVIAGRVLTCPSPDFSLFVPPNVQKTKPFSVPNLPLN 240

QY 241 TLSNRVPSLIKSMVSRDHQVQFQNGRVTLTGQGLTPTTSASQLCKIRGSVFHANG 300  
DB 241 TLSNRVPSLIKSMVSRDHQVQFQNGRVTLTGQGLTPTTSASQLCKIRGSVFHANG 300

QY 301 GNGYNLTDLGSPYHAFESPAPIGFDLGECDHMEASPTTQFNTGVDVIKQINVKQESAF 360  
DB 301 GNGYNLTDLGSPYHAFESPAPIGFDLGECDHMEASPTTQFNTGVDVIKQINVKQESAF 360

QY 361 APHLGTIOADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVPIPRYGSTITEAAQLAPP 420  
DB 361 APHLGTIOADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVPIPRYGSTITEAAQLAPP 420

QY 421 IYPGFGAIVFFMSDRPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480  
DB 421 IYPGFGAIVFFMSDRPIAHGTNGLSVPCCTIPOEFVTHFVNEQAPTRGEAALLHYLDPD 480

QY 481 THNLGFEKLYPEGFMTCPVNSSGTGPTLPINGVVFVSVSRFYQLKPVGTAGPACRL 540  
DB 481 THNLGFEKLYPEGFMTCPVNSSGTGPTLPINGVVFVSVSRFYQLKPVGTAGPACRL 540

## RESULT 5

QY1H12 PRELIMINARY; PRT; 545 AA.  
ID QY1H12  
AC QY1H12;

DT	01-DEC-2001 (TrEMBLrel. 19, Created)	DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	DE	Capsid protein.
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	OS	Human calicivirus HU/NLV/Birmingham/93/UK.
DE	Capsid.	OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OS	Human calicivirus NLV/VA98115/1998.	OC	Norovirus.
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;	OX	NCBI_TaxID=165509;
OC	Norovirus.	RN	[1]
OX	NCBI_TaxID=165509;	RP	SEQUENCE FROM N.A.
RN	[1]	RC	STRAIN=NLV/VA98115/1998;
RP	SEQUENCE FROM N.A.	RC	STRAIN=HU/NLV/Birmingham/93/UK;
RC	STRAIN=NLV/VA98115/1998;	RX	MEDLINE=20404883; PubMed=1094950;
RC	STRAIN=NLV/VA98115/1998;	RA	Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA	Fulton D., Morrow R., Matson D.O.;	RA	Clegg J.C., Chamberlain J., Brown D.W.G.;
RA	"Baculovirus expression and antigenic characterization of the capsid	RT	"Capsid protein diversity among 'Norwalk-like' viruses.";
RT	proteins of three Norwalk-like viruses.";	RL	Virus Genes 20:227-236(2000).
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.	EMBL	AU277612; CAB89093.1;
RL	EMBL; AY038598; AAK4673.1;	DR	InterPro; IPR004005; Calici_coat.
DR	InterPro; IPR004005; Calici_coat.	DR	InterPro; IPR008975; Viral_cap_coat.
DR	InterPro; IPR008975; Viral_cap_coat.	DR	Pfam; PF00915; Calici_coat; 1.
DR	Pfam; PF00915; Calici_coat; 1.	SQ	SEQUENCE 545 AA; 58295 MW; 258B0C7DEDB34DAC CRC64;
SQ	SEQUENCE 545 AA; 58295 MW; 258B0C7DEDB34DAC CRC64;		
Query Match	96.6%; Score 2798; DB 12; Length 545;	Query Match	94.8%; Score 2744; DB 12; Length 545;
Best Local Similarity	96.7%; Pred. No. 4.2e-224; Indels 0; Gaps 0;	Best Local Similarity	95.8%; Pred. No. 1.3e-219; Indels 0; Gaps 0;
Matches 527; Conservative 7; Mismatches 11; Indels 0; Gaps 0;		Matches 522; Conservative 7; Mismatches 16; Indels 0; Gaps 0;	
QY	1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60	QY	1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60
DB	1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60	DB	1 MMASKDAPTNMDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGQVNMIDPWIMNNY 60
QY	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSLHAQYNGWGNMVKVLLAGNAFTA 120	QY	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSLHAQYNGWGNMVKVLLAGNAFTA 120
DB	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSLHAQYNGWGNMVKVLLAGNAFTA 120	DB	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSLHAQYNGWGNMVKVLLAGNAFTA 120
QY	121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLFNHNDNAPTMR 180	QY	121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLFNHNDNAPTMR 180
DB	121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLFNHNDNAPTMR 180	DB	121 GKIIISCIIPGFAAQNISIAQATMFPHVIADVRVLEPIEVPLEDVRNVLFNHNDNAPTMR 180
QY	181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDPSFLFLVPPNVEQTKPFSVNPPLN 240	QY	181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDPSFLFLVPPNVEQTKPFSVNPPLN 240
DB	181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDPSFLFLVPPNVEQTKPFSVNPPLN 240	DB	181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDPSFLFLVPPNVEQTKPFSVNPPLN 240
QY	241 TLNSRVPSLTKSMVSRDHGMVQFQNGRVTLTGQQTTPTSASQLCKIRGSVFHANG 300	QY	241 TLNSRVPSLTKSMVSRDHGMVQFQNGRVTLTGQQTTPTSASQLCKIRGSVFHANG 300
DB	241 TLNSRVPSLTKSMVSRDHGMVQFQNGRVTLTGQQTTPTSASQLCKIRGSVFHANG 300	DB	241 TLNSRVPSLTKSMVSRDHGMVQFQNGRVTLTGQQTTPTSASQLCKIRGSVFHANG 300
QY	301 GNGYNLTDELGSPYHAFESPAPIGFPDLGECMDHMEASPTTQFTGDIKQINVKQESAF 360	QY	301 GNGYNLTDELGSPYHAFESPAPIGFPDLGECMDHMEASPTTQFTGDIKQINVKQESAF 360
DB	301 GNGYNLTDELGSPYHAFESPAPIGFPDLGECMDHMEASPTTQFTGDIKQINVKQESAF 360	DB	301 GNGYNLTDELGSPYHAFESPAPIGFPDLGECMDHMEASPTTQFTGDIKQINVKQESAF 360
QY	361 APHLGTTQADGLSDVSNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQLAPP 420	QY	361 APHLGTTQADGLSDVSNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQLAPP 420
DB	361 APHLGTTQADGLSDVSNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQLAPP 420	DB	361 APHLGTTQADGLSDVSNTNMIKLGWSPVSDGHRGVDVDPWIPRYGSTLTEAAQLAPP 420
QY	421 IYPGFGGEAIVFMSDFPIAHGNTGLSVPTCTIQEFVTHFVNEQAQTRGEAALLHYLDDP 480	QY	421 IYPGFGGEAIVFMSDFPIAHGNTGLSVPTCTIQEFVTHFVNEQAQTRGEAALLHYLDDP 480
DB	421 IYPGFGGEAIVFMSDFPIAHGNTGLSVPTCTIQEFVTHFVNEQAQTRGEAALLHYLDDP 480	DB	421 IYPGFGGEAIVFMSDFPIAHGNTGLSVPTCTIQEFVTHFVNEQAQTRGEAALLHYLDDP 480
QY	481 THRNLGFEKLYPEGFMTCPVNSSGTGQTLPIGNGVVFVSWSRFYQLKPVGTAGPACRL 540	QY	481 THRNLGFEKLYPEGFMTCPVNSSGTGQTLPIGNGVVFVSWSRFYQLKPVGTAGPACRL 540
DB	481 THRNLGFEKLYPEGFMTCPVNSSGTGQTLPIGNGVVFVSWSRFYQLKPVGTAGPACRL 540	DB	481 THRNLGFEKLYPEGFMTCPVNSSGTGQTLPIGNGVVFVSWSRFYQLKPVGTAGPACRL 540
QY	541 GIRRS 545	QY	541 GIRRS 545
DB	541 GIRRS 545	DB	541 GIRRS 545
RESULT 6		RESULT 7	
Q9IV45	PRELIMINARY; PRT; 545 AA.	Q66418	PRELIMINARY; PRT; 544 AA.
ID	Q9IV45	ID	Q66418
AC	Q9IV45;	AC	Q66418;
DT	01-OCT-2000 (TrEMBLrel. 15, Created)	DT	01-NOV-1996 (TrEMBLrel. 01, Created)
DT	01-OCT-2000 (TrEMBLrel. 15, Last sequence update)	DT	01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DE	Capsid protein.	DE	Capsid protein.

OS Desert Shield virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=33755;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSV395;  
RX MEDLINE=94174735; PubMed=8128635;  
RA Lew J.F., Kapikian A.Z., Jiang X., Estes M.K., Green K.Y.;  
RT "Molecular characterization and expression of the capsid protein of a  
RT Norwalk-like virus recovered from a Desert Shield troop with  
RT gastroenteritis";  
RL US94063772004Aug2325(W94)496  
DR EMBL; U04469; AAA16285.1; -  
DR InterPro; IPR004005; Calici.coat.  
DR Pfam; PF00915; Calici.coat; 1.  
SQ SEQUENCE 544 AA; 58606 MW; 07A4AE74BC581E8 CRC64;  
  
Query Match 90.0%; Score 2605.5; DB 12; Length 544;  
Best Local Similarity 89.2%; Pred. No. 4.4e-208;  
Matches 486; Conservative 24; Mismatches 34; Indels 1; Gaps 1;  
  
Qy 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNY 60  
Db 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNSY 60  
  
Qy 61 VOAPQGETTISPNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120  
Db 61 VOAPQGETTISPNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120  
  
Qy 121 GKIIISCIPTPGFAAQNISIAQATWPHVIAADVRLIEVPLEDVRNVLPHNNDNAPTMR 180  
Db 121 GKIIISCIPTPGFAAQNISIAQATWPHVIAADVRLIEVPLEDVRNVLPHNNDSSPTMR 180  
  
Qy 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCTSPDFSLFLVPPNVQKTKPFSVPNLPLN 240  
Db 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCTSPDFSLFLVPPNVQKTKPFSVPNLPLN 240  
  
Qy 241 TLNSRVPSLTKSMVSRDHQOMQVQFNGRVTLQGLQGTTPTSASQCKIRGVSFHANG 300  
Db 241 TLNSRVPSLTKSMVSRDHQOMQVQFNGRVTLQGLQGTTPTSASQCKIRGVSFHANG 300  
  
Qy 301 GNGYNLTLDGSPYHAFESPAPIGPFDLGECDWMEASPTTFQNTGDIKQINVKQESAF 360  
Db 301 GNGYNLTLDGSPYHAFESPAPIGPFDLGECDWMEASPTTFQNTGDIKQINVKQESAF 360  
  
Qy 361 APHLGTVQADGLSDVSVNTNMIAGLWVSPVSDGHRGVDVDPWVIPRYGSTLTTEAAQLAPP 420  
Db 361 APHLGTVQADGLSDVSVNTNMIAGLWVSPVSDGHRGVDVDPWVIPRYGSTLTTEAAQLAPP 420  
  
Qy 421 IYPGFGGAIYFFMSDFPIAHGTNGLSVPCCTIPQEFVTHFVNEQAPTRGEAALLHYLDPD 480  
Db 421 IYPGFGGAIYFFMSDFPIAHGTNGLSVPCCTIPQEFVTHFVNEQAPTRGEAALLHYLDPD 480  
  
Qy 481 THRLNLFGEFKLYPEGFMTCVPNSSGTGPTLPIGVFVSVMSVSRFYQLKPVGTAGPACRL 540  
Db 481 THRLNLFGEFKLYPEGFMTCVPNSSGTGPTLPIGVFVSVMSVSRFYQLKPVGTAGPACRL 540  
  
Qy 541 GIRRS 545  
Db 540 GIRRS 544  
  
RESULT 8  
Q8BCA3 ID Q8BCA3 PRELIMINARY; PRT; 543 AA.  
AC Q8BCA3;  
DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid.  
OS Human calicivirus NLV/Boxer/2001/US.

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=207658;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NLV/Boxer/2001/US;  
RX Farkas T., Thornton S.A., Zhong W., Jiang X.;  
RA "Viral gastroenteritis outbreaks on US Navy Ships";  
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.  
RE EMBL; AF538679; AAN15140.1; -  
DR InterPro; IPR004005; Calici.coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calici.coat; 1.  
SQ SEQUENCE 543 AA; 58783 MW; 9CFD55A1858E0F62 CRC64;  
  
Query Match 73.9%; Score 2139; DB 12; Length 543;  
Best Local Similarity 72.1%; Pred. No. 3e-169;  
Matches 397; Conservative 74; Mismatches 66; Indels 14; Gaps 7;  
  
Qy 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNY 60  
Db 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGVNNIDPWIMNF 60  
  
Qy 61 VOAPQGETTISPNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120  
Db 61 VOAPQGETTISPNTPGDILFDLQGLPHNLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120  
  
Qy 121 GKIIISCIPTPGFAAQNISIAQATWPHVIAADVRLIEVPLEDVRNVLPHNNDNAPTMR 180  
Db 121 GKIIISCIPTPGFAAQNISIAQATWPHVIAADVRLIEVPLEDVRNVLPHNNDNAPTMR 180  
  
Qy 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCTSPDFSLFLVPPNVQKTKPFSVPNLPLN 240  
Db 181 LVCMLYTPLRASGSSGTDPPFVIAGRVLTCTSPDFSLFLVPPNVQKTKPFSVPNLPLN 240  
  
Qy 241 TLNSRVPSLTKSMVSRDHQOMQVQFNGRVTLQGLQGTTPTSASQCKIRGVSFHANG 300  
Db 241 TLNSRVPSLTKSMVSRDHQOMQVQFNGRVTLQGLQGTTPTSASQCKIRGVSFHANG 300  
  
Qy 301 GNGYNLTLDGSPYHAFESPAPIGPFDLGECDWMEASPTTFQNTGDIKQINVKQESAF 356  
Db 301 GNGYNLTLDGSPYHAFESPAPIGPFDLGECDWMEASPTTFQNTGDIKQINVKQESAF 356  
  
Qy 357 ESAPFAPLGTITQADGLSDVSVNTNMIAGLWVSPVSDGHRGVDVDPWVIPRYGSTLTTEA 414  
Db 357 ESAPFAPLGTITQADGLSDVSVNTNMIAGLWVSPVSDGHRGVDVDPWVIPRYGSTLTTEA 414  
  
Qy 415 AQLAPPYPPGFGGAIYFFMSDFPIAHGTNGLSVPCCTIPQEFVTHFVNEQAPTRGEAALL 474  
Db 415 AQLAPPYPPGFGGAIYFFMSDFPIAHGTNGLSVPCCTIPQEFVTHFVNEQAPTRGEAALL 474  
  
Qy 475 HYLPDPTHRLNLFGEFKLYPEGFMTCVPNSSGTGPTLPIGVFVSVMSVSRFYQLKPVGTA 534  
Db 475 HYLPDPTHRLNLFGEFKLYPEGFMTCVPNSSGTGPTLPIGVFVSVMSVSRFYQLKPVGTA 534  
  
Qy 535 GPACRLGIRRS 545  
Db 533 GAARRLGIRRS 543  
  
RESULT 9  
Q9IV48 ID Q9IV48 PRELIMINARY; PRT; 539 AA.  
AC Q9IV48;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus HU/NLV/Winchester/94/UK.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=122913;  
RN [1]

```
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Winchester/94/UK;
RX MEDLINE=20404883; PubMed=1094950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL; AJ277609; CAB89090.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 539 AA; 58120 MW; 14E612D7A576395D CRC64;

Query Match 73.3%; Score 2124; DB 12; Length 539;
Best Local Similarity 72.7%; Pred. No. 5, 2e-168;
Matches 397; Conservative 56; Mismatches 83; Indels 10; Gaps 6;

QY 1 MMASKDAPTNNMDGTSGAGQLVPEANTAEPISEMEPVAGAATAAATAGQVNMIDPMINNY 60
DB 1 MMASKDVPSNMDXSGAGQLVPEVNAEPLPLEVVGAAATAAATAGQVNLIDPMINNF 60

QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQYNGWGMKVKVLLAGNAFTA 120

QY 121 GKIISCIPPGFAAQNISIAQATMPFHVIADRVLEIEVPLEDVRNVLFNHNDVPTMR 180
DB 121 GKIICCVPPGASQNISIGQATMPFHVIADRVLEIEIPLDDVRNVLFTNENRPTMR 180

QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFVSVPNLPLN 240
DB 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFVSVPNLPLN 240

QY 181 LLCMLYTPLRAGASSSGTDPFVIAGRVLTCPSPDFNLFVLPVPSVEQTRQLTVENIPLN 240
DB 181 LLCMLYTPLRAGASSSGTDPFVIAGRVLTCPSPDFNLFVLPVPSVEQTRQLTVENIPLN 240

QY 241 TLSNSRVPSLIKSMVSRDHQGMVQFNGRVTLTGQGLTTPTSASQLCKIRGSVFHANG 300
DB 241 NLANSRVPMINKVTVSTQNVQVQFNGRCILEGQLLTTPVSASQVARIKGVSTAS 300

QY 301 GNGYNLTDLGSPYHAFESPAPTFGPDLCEDWHM--EASPTQFNTGVVIQINVKQES 358
DB 301 GKGLNLTDLGTPYHAFESPAPTFGPDLCEDWHM--EASPTQFNTGVVIQINVKQES 358

QY 359 AFPHLGTIOADGLSDVSVNTNMIKLGWVSPVSDGHRGVDVDPWVPIPRYGSSTLTAQAQLA 418
DB 359 PICTXGLSIEF--TSDDQDPAQGLQGLTAWVSPSTSGAR--VDPWKIPSYGSTVTESTHLA 414

QY 419 PPIYPGFGAIVFMSDFPIAHGTNGLSVPCITPQEFVTHFVNEQAPRGEAALLHYLD 478
DB 415 PLIFPPGFGAIVFMSDFPIVSG-NTAQIPCTLPQEFVSHFVEQQAPIRGEAALLHYVD 473

QY 479 PDTHRNLFGEFLKLYPEGFWTCVPSNGGTPQTLPIGVFVSVWSRYQKPKVGTAGPAC 538
DB 474 PDTHRNLFGEFLKLYPDGFITCVNTGG-GPQNLPINGVVFVSSWSRYQKPKVGTAGPAR 532

QY 539 RLGIIR 544
DB 533 RLGVRR 538

RESULT 10
Q8XJJO PRELIMINARY; PRT; 543 AA.
AC Q8XJJO;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SzUG1;

RA Katayama K., Kojima S.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RA SEQUENCE FROM N.A.
RC STRAIN=SzUG1;
RX MEDLINE=22192455; PubMed=12202225;
RA Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,
RA Gojobori T., Takeda N.;
RA "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
RT Viruses.";
RL Virology 299:225-239(2002).
DR EMBL; AB039774; BAC11813.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 543 AA; 58761 MW; E0DE7489A9C3488C CRC64;

Query Match 69.7%; Score 2017.5; DB 12; Length 543;
Best Local Similarity 70.1%; Pred. No. 3.9e-159;
Matches 385; Conservative 58; Mismatches 95; Indels 11; Gaps 7;

QY 1 MMASKDAPTNNMDGTSGAGQLVPEANTAEPISEMEPVAGAATAAATAGQVNMIDPMINNY 60
DB 1 MMASKDAPTSGADGANGAGQLVPEVNAEPLPLDPVAGASTALATAGQVNMIDPMINNF 60

QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQYNGWGMKVKVLLAGNAFTA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQYNGWGMKVKVLLAGNAFTA 120

QY 121 GKIISCIPPGFAAQNISIAQATMPFHVIADRVLEIEVPLEDVRNVLFNHNDVPTMR 180
DB 121 GKIICCVPPGFGOSRTLSIAQATLPFHVIADRVLEIEIPLDVRNVLFNHNDVPTMR 180

QY 181 LVCMLYTPLRASGSSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFVSVPNLPLN 240
DB 181 LLCMLYTPLRGSGSGTDAFVAGRVLTCPSPDFNLFVLPVPTVBQKTRPFSVPNIPLQ 240

QY 241 TLSNSRVPSLIKSMVSRDHQGMVQFNGRVTLTGQGLTTPTSASQLCKIRGSVFHANG 300
DB 241 NLSNSRVPSLIQSMVLSNDAQTQVQFNGRCITDGHLLGTTTPVSASQLMKFRGV--TSG 298

QY 301 GNGYNLTDLGSPYHAFESPAPTFGPDLCEDWHM--EASPTQFNTGVVIQINVKQES 359
DB 299 SKVNLNLTDLGSPPLAEAPAPAGPDLGKCDWHVMSLYQDNNHADPIVLHAIEPNSSS 358

QY 360 PAPHLGTIOADGLSDVSVNTNMIKLGWVSPVSDGH--RGVDVDPWVPIPRYGSSTLTAQAQL 417
DB 359 FVPHLGSVSFK--ENNVNAGDVYVCTIQYTSPPSDPHNPNADVDFWSIPDYGSLAEASQL 416

QY 418 APPYPPGFGAIVFMSDFPIAHGTNGLSVPCITPQEFVTHFVNEQAPRGEAALLHY 476
DB 417 APVYPPGFGAIVFMSDFPIAHGTNGLSVPCITPQEFVTHFVNEQAPRGEAALLHY 476

QY 477 LDPDTHRNLFGEFLKLYPEGFWTCVPSNGGTPQTLPIGVFVSVWSRYQKPKVGTAGP 536
DB 477 VDPDTHRNLFGEFLKLYPEGFWTCVPSNGGTPQTLPIGVFVSVWSRYQKPKVGTAGP 534

QY 537 A-CRIGIRR 544
DB 535 ARGRLGIIR 543

RESULT 11
Q9DU46 PRELIMINARY; PRT; 544 AA.
AC Q9DU46;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
```



```
OC Norovirus.
OX NCBI_TaxID=99565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Chiba 407/1987/Jp;
RA Someya Y., Takeda N., Miyamura T.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Chiba 407/1987/Jp;
RA MEDLINE=20569531; PubMed=11118371;
RX Someya Y., Takeda N., Miyamura T.;
RT "Complete nucleotide sequence of the Chiba virus genome and functional
RT expression of the 3C-like protease in Escherichia coli.";
RL Virology 278:490-500(2000).
DR EMBL; AB042808; BAB18267.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58351 MW; 1DB7A4E81FBB246F CRC64;

Query Match 68.3%; Score 1979; DB 12; Length 544;
Best Local Similarity 66.8%; Pred. No. 6.2e-156;
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;

Qy 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMIDPWIMNYY 60
Db 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMIDPWIMNYY 60
Qy 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
Qy 121 GKIISICPPGFAAQNISIAQATMFPHVIADRVLEPVEPLEVDRVNLPHNNDAPTMR 180
Db 121 GKIISICPPGFAAQNISIAQATMFPHVIADRVLEPVEPLEVDRVNLPHNNDAPTMR 180
Qy 181 LVCMLYTLPLRASGSSGTDPPVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSVPNPLN 240
Db 181 LVCMLYTLPLRASGSSGTDPPVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSVPNPLN 240
Qy 241 TLSNRVPSLIKMMVSRDHQMVQFQNGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300
Db 241 TLSNRVPSLIKMMVSRDHQMVQFQNGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300
Qy 301 GNGVNLTELDGSPVHAFESAPIGFDPDLGECDHMEASPT--PTQFNTGDVVKQINVAQES 358
Db 301 GNGVNLTELDGSPVHAFESAPIGFDPDLGECDHMEASPT--PTQFNTGDVVKQINVAQES 358
Qy 359 AFAPHLGTIOADGLSDVSVNTNMIAKLGWSPVSDGHRGDVDPWVI PRYGSTLTAAQA 418
Db 359 AFAPHLGTIOADGLSDVSVNTNMIAKLGWSPVSDGHRGDVDPWVI PRYGSTLTAAQA 418
Qy 419 PPIYPGFGBAIVFFMSDFPIAHGTNGLS----VPCTIPQEFVTHFVNEQAPTGEAALL 474
Db 419 PPIYPGFGBAIVFFMSDFPIAHGTNGLS----VPCTIPQEFVTHFVNEQAPTGEAALL 474
Qy 475 HYLDPDTHRLNGLGEFKLYPEGFMTCVPNSSGTPQTLPINGVVFVSVWSRFPYQLKPVGTA 534
Db 475 HYLDPDTHRLNGLGEFKLYPEGFMTCVPNSSGTPQTLPINGVVFVSVWSRFPYQLKPVGTA 534
Qy 535 GPA-CRLGIRR 544
Db 535 GPA-CRLGIRR 544

RESULT 12
Q8JW44
ID Q8JW44 PRELIMINARY; PRT; 546 AA.
AC Q8JW44;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Capsid protein.
GN CAPSID.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Funabashi258;
RA Katayama K., Takeda N., Natori K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Funabashi258;
RA Natori K., Takeda N.;
RT "Genetic and antigenic relationship among Norwalk-like viruses.";
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB078335; BAC05516.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 546 AA; 58834 MW; 1965F054E2C48186 CRC64;

Query Match 68.3%; Score 1977.5; DB 12; Length 546;
Best Local Similarity 66.6%; Pred. No. 8.3e-156;
Matches 367; Conservative 76; Mismatches 95; Indels 13; Gaps 7;

Qy 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMIDPWIMNYY 60
Db 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAAGOVNMIDPWIMNYY 60
Qy 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLAQMYNGWGNMKVLLAGNAFTA 120
Qy 121 GKIISICPPGFAAQNISIAQATMFPHVIADRVLEPVEPLEVDRVNLPHNNDAPTMR 180
Db 121 GKIISICPPGFAAQNISIAQATMFPHVIADRVLEPVEPLEVDRVNLPHNNDAPTMR 180
Qy 181 LVCMLYTLPLRASGSSGTDPPVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSVPNPLN 240
Db 181 LVCMLYTLPLRASGSSGTDPPVIAGRVLTCPSPDFSEFLVPPNVEQTKPFSVPNPLN 240
Qy 241 TLSNRVPSLIKMMVSRDHQMVQFQNGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300
Db 241 TLSNRVPSLIKMMVSRDHQMVQFQNGRVTLTGQLOGTTPTTSASOLCKIRGSVFHANG 300
Qy 301 GNGVNLTELDGSPVHAFESAPIGFDPDLGECDHMEASPT--TQNTGDVVKQINVAQES-S 358
Db 301 GNGVNLTELDGSPVHAFESAPIGFDPDLGECDHMEASPT--TQNTGDVVKQINVAQES-S 358
Qy 359 AFAPHLGTIOADGLSDVSVNTNMIAKLGWSPVSDGHRGDVDPWVI PRYGSTLTAAQA 418
Db 359 AFAPHLGTIOADGLSDVSVNTNMIAKLGWSPVSDGHRGDVDPWVI PRYGSTLTAAQA 418
Qy 419 PPIYPGFGBAIVFFMSDFPIAHGTNGLS----VPCTIPQEFVTHFVNEQAPTGEAALL 474
Db 419 PPIYPGFGBAIVFFMSDFPIAHGTNGLS----VPCTIPQEFVTHFVNEQAPTGEAALL 474
Qy 475 HYLDPDTHRLNGLGEFKLYPEGFMTCVPNSSGTPQTLPINGVVFVSVWSRFPYQLKPVGTA 534
Db 475 HYLDPDTHRLNGLGEFKLYPEGFMTCVPNSSGTPQTLPINGVVFVSVWSRFPYQLKPVGTA 534
Qy 535 GPA-CRLGIRR 544
Db 535 GPA-CRLGIRR 544

RESULT 13
Q91185
ID Q91185 PRELIMINARY; PRT; 544 AA.
AC Q91185;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
```

DT	01-DEC-2001 (TrEMBLrel. 19, Last sequence update)	OS	Chiba virus.
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
DE	Capsid protein.	OC	Norovirus.
OS	Human calicivirus HU/NLV/Koblentz/433/2000/DE.	OX	NCBI_TaxID=99565;
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;	RN	[1]
OC	Norovirus.	RP	SEQUENCE FROM N.A.
OX	NCBI_TaxID=165252;	RC	STRAIN=Chiba 407;
RN	[1]	RA	Natori K., Suzuki K., Yamakawa Y., Tatsumi M., Sakae K., Kobayashi S.,
RP	SEQUENCE FROM N.A.	RA	Shinozaki K., Ishiko H., Miyamura T., Takeda N.;
RC	STRAIN=HU/NLV/Koblentz 433/2000/DE;	RT	"Expression and self-assembly of capsid proteins of the Chiba virus, a
RA	Kuenkel U., Schreier E.;	RT	genetically distinct Norwalk-like virus.";
RT	"Molecular epidemiology of outbreaks of gastroenteritis associated	RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.
RT	with Norwalk-like viruses in Germany.";	DR	EMBL; AB022679; BAA82106.1; -
RL	Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.	DR	InterPro; IPR004005; Calici.coat.
DR	EMBL; AF394960; AAK72048.1; -	DR	InterPro; IPR008975; Viral_Cap_coat.
DR	InterPro; IPR004005; Calici.coat.	DR	Pfam; PF00915; Calici.coat.1.
DR	InterPro; IPR008975; Viral_Cap_coat.	SQ	SEQUENCE 544 AA; 58323 MW; ABDLC1FC4F93D872 CRC64;
DR	Pfam; PF00915; Calici.coat; 1.		
SQ	SEQUENCE 544 AA; 58222 MW; 1B39A595B733A6A8 CRC64;		
Query Match 68.2%; Score 1975; DB 12; Length 544;			
Best Local Similarity 66.8%; Pred. No. 1.3e-155;			
Matches 368; Conservative 68; Mismatches 101; Indels 14; Gaps 6;			
Qy	1 MMASKDAPTNDGTSAGOLVPEANTAEPISEMPVAGATAAATAGOVNIDPWIMNY 60	Qy	1 MMASKDAPTNDGTSAGOLVPEANTAEPISEMPVAGATAAATAGOVNIDPWIMNY 60
Db	1 MMASKDAPTNDGTSAGOLVPEANTAEPISEMPVAGATAAATAGOVNIDPWIMNY 60	Db	1 MMASKDAPTNDGTSAGOLVPEANTAEPISEMPVAGATAAATAGOVNIDPWIMNY 60
Qy	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMVKVLLAGNAFTA 120	Qy	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMVKVLLAGNAFTA 120
Db	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMVKVLLAGNAFTA 120	Db	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGWGNMVKVLLAGNAFTA 120
Qy	121 GKIIISCIIPGFAAQNISIAQATMFPHVJADVRVLEPIEVLDPEDVRNVLPHNNDNAPTMR 180	Qy	121 GKIIISCIIPGFAAQNISIAQATMFPHVJADVRVLEPIEVLDPEDVRNVLPHNNDNAPTMR 180
Db	121 GKIIISCIIPGFAAQNISIAQATMFPHVJADVRVLEPIEVLDPEDVRNVLPHNNDNAPTMR 180	Db	121 GKIIISCIIPGFAAQNISIAQATMFPHVJADVRVLEPIEVLDPEDVRNVLPHNNDNAPTMR 180
Qy	181 LVCMLYPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQKTKPFSVNPPLN 240	Qy	181 LVCMLYPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQKTKPFSVNPPLN 240
Db	181 LVCMLYPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQKTKPFSVNPPLN 240	Db	181 LVCMLYPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVEQKTKPFSVNPPLN 240
Qy	241 TLNSRVPSLIKSMVSRDHGQVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHANG 300	Qy	241 TLNSRVPSLIKSMVSRDHGQVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHANG 300
Db	241 TLNSRVPSLIKSMVSRDHGQVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHANG 300	Db	241 TLNSRVPSLIKSMVSRDHGQVQFQNGRVTLDGQLQGTTPTSASQLCKIRGSVFHANG 300
Qy	301 GNGYNLTDLGSPYHAFESPAPIGFDLGECDWHMEAS--PTQFNTGDIKQINVKQES 358	Qy	301 GNGYNLTDLGSPYHAFESPAPIGFDLGECDWHMEAS--PTQFNTGDIKQINVKQES 358
Db	301 GNGYNLTDLGSPYHAFESPAPIGFDLGECDWHMEAS--PTQFNTGDIKQINVKQES 358	Db	301 GNGYNLTDLGSPYHAFESPAPIGFDLGECDWHMEAS--PTQFNTGDIKQINVKQES 358
Qy	358 QKVLNLTDLGSPYHAFESPAPIGFDLGECDWHMEAS--PTQFNTGDIKQINVKQES 358	Qy	358 QKVLNLTDLGSPYHAFESPAPIGFDLGECDWHMEAS--PTQFNTGDIKQINVKQES 358
Db	358 QKVLNLTDLGSPYHAFESPAPIGFDLGECDWHMEAS--PTQFNTGDIKQINVKQES 358	Db	358 QKVLNLTDLGSPYHAFESPAPIGFDLGECDWHMEAS--PTQFNTGDIKQINVKQES 358
Qy	359 AFAPHLGTIQADGLSDVSVNTNMIAGLVSPVSDGHRGVDVDPWVPIRYGSLTTEAAQLA 418	Qy	359 AFAPHLGTIQADGLSDVSVNTNMIAGLVSPVSDGHRGVDVDPWVPIRYGSLTTEAAQLA 418
Db	359 AFAPHLGTIQADGLSDVSVNTNMIAGLVSPVSDGHRGVDVDPWVPIRYGSLTTEAAQLA 418	Db	359 AFAPHLGTIQADGLSDVSVNTNMIAGLVSPVSDGHRGVDVDPWVPIRYGSLTTEAAQLA 418
Qy	419 PPIYPPGFGGAIVFMSDFPIAHGTNGLS----VPCTIPOEFVTHFVNEQAPRGEAALL 474	Qy	419 PPIYPPGFGGAIVFMSDFPIAHGTNGLS----VPCTIPOEFVTHFVNEQAPRGEAALL 474
Db	419 PPIYPPGFGGAIVFMSDFPIAHGTNGLS----VPCTIPOEFVTHFVNEQAPRGEAALL 474	Db	419 PPIYPPGFGGAIVFMSDFPIAHGTNGLS----VPCTIPOEFVTHFVNEQAPRGEAALL 474
Qy	473 PAVYPPGFEVIVVFMAISIP---GPNQSGSPNLVPCLLPQEIYTHFISQAPIQGEAALL 473	Qy	473 PAVYPPGFEVIVVFMAISIP---GPNQSGSPNLVPCLLPQEIYTHFISQAPIQGEAALL 473
Db	473 PAVYPPGFEVIVVFMAISIP---GPNQSGSPNLVPCLLPQEIYTHFISQAPIQGEAALL 473	Db	473 PAVYPPGFEVIVVFMAISIP---GPNQSGSPNLVPCLLPQEIYTHFISQAPIQGEAALL 473
Qy	475 HYLPDPTNRNLGEFKLYPEGFMTCPVNSSTGTPQLPFGVVFVSWSRFYQLKPVGTA 534	Qy	475 HYLPDPTNRNLGEFKLYPEGFMTCPVNSSTGTPQLPFGVVFVSWSRFYQLKPVGTA 534
Db	475 HYLPDPTNRNLGEFKLYPEGFMTCPVNSSTGTPQLPFGVVFVSWSRFYQLKPVGTA 534	Db	475 HYLPDPTNRNLGEFKLYPEGFMTCPVNSSTGTPQLPFGVVFVSWSRFYQLKPVGTA 534
Qy	535 GPA-CRLGIRR 544	Qy	535 GPA-CRLGIRR 544
Db	535 GPA-CRLGIRR 544	Db	535 GPA-CRLGIRR 544
Qy	544 GPARGRLGVR 544	Qy	544 GPARGRLGVR 544
Db	544 GPARGRLGVR 544	Db	544 GPARGRLGVR 544
RESULT 14			
Q9QTE7			
ID	Q9QTE7	PRELIMINARY;	PRT; 544 AA.
AC	Q9QTE7;		
DT	01-MAY-2000 (TrEMBLrel. 13, Created)		
DT	01-MAY-2000 (TrEMBLrel. 13, Last sequence update)		
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)		
DE	Capsid protein.		

```
OC Norovirus.
OX NCBI_TaxID=11983;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Norwalk-like;
RX MEDLINE=20254531; PubMed=10795514;
RA Schreier E., Doering F., Kuenkel U.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RL with small round structured viruses in Germany in 1997/98.";
RN Arch. Virol. 145:443-453(2000).
RP [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Norwalk-like;
RA Schreier E., Doering F., Kuenkel U.;
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF093797; AAC64603.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 540 AA; 58018 MW; 647843487A654CBE CRC64;

Query Match      68.1%; Score 1972.5; DB 12; Length 540;
Best Local Similarity 67.7%; Pred. No. 2.1e-155;
Matches 377; Conservative 60; Mismatches 89; Indels 31; Gaps 7;

Qy 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISEBPVAGAATAAATAGQVNNIDPWIMNY 60
Db 1 MMASKDAPSPDGASGAGQLVPEVTADQISMDPVAGASTAVATAGQVNNIDPWIFNF 60

Qy 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQYNGVGNMKVLLAGNAFTA 120
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLAHLNQMYNGVGNMRVRIILAGNAFTA 120

Qy 121 GKIIISCTIPGFAAQNTISAOATWPHVIAADVRLPEIEVLEPVEDVRNVLPHNNDN-APT 179
Db 121 GKVIICVPPGFDARILHIAQTULPHLIADVRTLEPVELEPVEDVRNVLHNSSQPOPTM 180

Qy 180 RLVCMLYTLPRASGSSGTDTPFVIAGRVLTCPSPDFSLFLVPPNVEQTKPFVSPNPL 239
Db 181 RLVMALYTLPLRTGGSGGTDFAFVAGRVLTCPAPDFSLFLVPPSVEQKTRVFSVPNIPL 240

Qy 240 NTLNSRVPSLIKSMVSRDHQWVQFQNGRVTLDGQLQGTTPTSASOLCKIRGSVFHAN 299
Db 241 KDLNSRVPTLIQGMFVSPDVNQSVQFQNGRCQIDGQLQGTTPVLSQLCKIRGT--SS 298

Qy 300 GGNGYNLTLDGSPVHAFESAPIGFDPDLGECDDHM-----EASPTTQFNTGQVIK 350
Db 299 NTRVLNLSEVDGTPFVPLESPAPVGFDPDGGCDWHVGTTFEARDQDPSONVTFATND--- 355

Qy 351 QINVKQESAFAPHLGTI---QADGLSDVSVNTNMIAKLGWVSPVSDGHRGDVDPWVIPRY 407
Db 356 -----SSFVPYLGSIAPHNGDGFH---SGDIIGSLDWISAPSDGSALDV--WSIPKY 402

Qy 408 GSTUTEAAQLAPPYIPGFGGAIIVFFMSDPPFIANGTNGLSVPCTIPQEFVTHFVNEOAPT 467
Db 403 GSSLPDVTHLAPAVFPFGFGEVILYFHSKFGSGPTDKLRVPCLIPQEFITHFCNEQAPI 462

Qy 468 RGEAALLHYLDPDTHRNLEPKLYPEGFMTCVPNSSGTGTOTLPINGVFVSVWSRFYQ 527
Db 463 AGEAALLHYVDPTGRNLEPKLYPDGFMTCVPNLSISGGPOTLPINGVFVSVWSRFYQ 522

Qy 528 LKPVGTAGPACRLGIRR 544
Db 523 LKPVGTASAARRLGIRR 539
```

Search completed: June 1, 2004, 13:53:09  
Job time : 33.7307 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 45.5302 Seconds  
(without alignments)  
3289.030 Million cell updates/sec

Title: US-09-926-799-2  
Perfect score: 2819  
Sequence: 1 MMASKDATSSVDGASGAGQ.....YQLKPVCTASSARGRLGLRR 530

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2819	100.0	530	4	AAB49701 Small rou
2	2793	99.1	530	2	AAR50972 Norwalk v
3	2793	99.1	530	7	ADC72176 Norwalk v
4	2713	96.2	530	2	AAR57091 Small rou
5	2093	74.2	544	4	AAB49703 Small rou
6	2029.5	72.0	546	4	AAB49702 Small rou
7	1938	68.7	545	4	AAB49700 Small rou
8	1232	43.7	540	4	AAB49706 Small rou
9	1209	42.9	548	4	AAB49705 Small rou
10	1200.5	42.6	535	4	AAB49707 Small rou
11	1198	42.1	548	5	AAB49707 Small rou
12	1181	41.9	542	4	AAB49708 Small rou
13	1148.5	40.7	539	4	AAB49704 Small rou
14	1131	40.1	550	4	AAB49709 Small rou
15	1005.5	35.7	541	4	AAB49710 Small rou
16	326	11.6	579	2	AAW08143 RHDV caps
17	244.5	8.7	668	4	AAB67462 Amino aci
18	240.5	8.5	668	2	AAR10686 Feline ca
19	240.5	8.5	668	4	AAE04304 Feline ca
20	239	8.5	623	4	AAB47043 Feline ca
21	238	8.4	669	4	AAB67461 Amino aci
22	237	8.4	547	4	AAW50108 Feline ca
23	237	8.4	671	4	AAW50107 Feline ca
24	236	8.4	623	4	AAB47044 Feline ca
25	225	8.0	622	4	AAB47045 Feline ca

26	136	4.8	40	5	AAU91273	Aau91273 Norwalk v
27	134.5	4.8	2164	1	AAP81045	Aap81045 Sequence
28	132.5	4.7	2164	1	AAP80131	Aap80131 Peptides
29	132	4.7	2206	2	AAR22210	Aar22210 True type
30	130	4.6	40	5	AAU91274	Aau91274 Norwalk v
31	127	4.5	517	2	AAW17975	Aaw17975 Coprinus
32	126.5	4.5	1049	6	ABU21976	Abu21976 Protein e
33	126	4.5	2209	1	AAP20037	Aap20037 Sequence
34	126	4.5	6310	6	ABU39869	Abu39869 Protein e
35	124.5	4.4	3892	6	ADA34216	Ada34216 Acinetoba
36	123.5	4.4	2183	3	AAB03533	Aab03533 Murine fa
37	123.5	4.4	2639	4	ABG15016	Abg15016 Novel hum
38	121.5	4.3	749	6	ABU27823	Abu27823 Protein e
39	119.5	4.2	3716	6	ABM15900	Abm15900 Mycobacte
40	119	4.2	1121	4	ABB59969	Abb59969 Drosophila
41	118.5	4.2	8360	6	AAE35499	Aae35499 Streptomy
42	118	4.2	918	2	AAW88422	Aaw88422 Chlamydia
43	117.5	4.2	1459	6	ABM15895	Abm15895 Mycobacte
44	117.5	4.2	1459	7	ADB80091	Adb80091 Mycobacte
45	116	4.1	2468	6	ABU38411	Abu38411 Protein e

ALIGNMENTS

RESULT 1  
AAB49701  
ID AAB49701 standard; protein; 530 AA.

AC AAB49701;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 2.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX (DENK-) DENKA SEIKEN KK.

Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
WPI: 2001-080848/09.  
DR N-PSDB; AAF29142.  
PT Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.  
PS Claim 1; Page 42-45; 84pp; Japanese.

This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

Sequence 530 AA;  
Query Match 100.0%; Score 2819; DB 4; Length 530;  
Best Local Similarity 100.0%; Pred. No. 1.5e-247;

Wed Jun 2 09:13:26 2004

us-09-926-799-2.rag

Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMASKDATSSVDGAGAGQVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINF 60  
DB 1 MMASKDATSSVDGAGAGQVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINF 60

QY 61 VOAQGEFTISPNNTPGGVLFDLSLGLPHLNPFLHLSQMYNGWGNVRVIMLAGNAFTA 120  
DB 61 VOAQGEFTISPNNTPGGVLFDLSLGLPHLNPFLHLSQMYNGWGNVRVIMLAGNAFTA 120

QY 121 GKIIIVSCIPPGFSGSHNLTTIAQATLFPFHVIADVRTLDPIEVLPLEDVNRVLFHNDRNQTM 180  
DB 121 GKIIIVSCIPPGFSGSHNLTTIAQATLFPFHVIADVRTLDPIEVLPLEDVNRVLFHNDRNQTM 180

QY 181 RLVCMLYTLPLRTGGTGGDSFVAGRVMTCPSPDFNLFVPTVEQKTRPFTLPNPLSS 240  
DB 181 RLVCMLYTLPLRTGGTGGDSFVAGRVMTCPSPDFNLFVPTVEQKTRPFTLPNPLSS 240

QY 241 LNSRAPLPISGMISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300  
DB 241 LNSRAPLPISGMISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300

QY 301 LTLDGTPFPFPGAPIGFDPDLGGCDWHINMTQFGHSSQTYDVTDPDTFVPHLGSIQ 360  
DB 301 LTLDGTPFPFPGAPIGFDPDLGGCDWHINMTQFGHSSQTYDVTDPDTFVPHLGSIQ 360

QY 361 ANGIGSNYIGVLSWVSPSPHSPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420  
DB 361 ANGIGSNYIGVLSWVSPSPHSPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420

QY 421 FMSKIPGAGYSLPCLLPQEIYISHLASEQAPTVEAALLHYVDPDGTGRTLGEFKAYPDGF 480  
DB 421 FMSKIPGAGYSLPCLLPQEIYISHLASEQAPTVEAALLHYVDPDGTGRTLGEFKAYPDGF 480

QY 481 LTCVPNGASSGQQQLPINGVFVFSWVSRRFYQLKPVGTASSARGRLGLRR 530  
DB 481 LTCVPNGASSGQQQLPINGVFVFSWVSRRFYQLKPVGTASSARGRLGLRR 530

RESULT 2  
AAR50972  
ID AAR50972 standard; protein; 530 AA.  
XX AC AAR50972;  
XX DT 16-OCT-2003 (revised)  
XX DT 25-MAR-2003 (revised)  
XX DT 05-OCT-1994 (first entry)  
XX DE Norwalk virus strain 8FIIa protein (encoded by ORF2).  
XX KW Norwalk virus; pathogen; acute gastroenteritis; food poisoning;  
XX KW seafood contamination; diagnostic assay; calcivirus; small round virus.  
XX OS Norwalk virus; (strain 8FIIa).  
XX PN **W09405700-A2**.  
XX PD 17-MAR-1994.  
XX PF 07-SEP-1993; 93WO-US008447.  
XX PR 07-SEP-1992; 92US-00941365.  
XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
XX PI **Natson DO, Estes MK, Jiang X, Graham DY;**  
XX DR WPI; 1994-101125/12.  
XX DR N-PSDB; AAQ56826.  
XX PT DNA from Norwalk and related viruses - used for preparing prods. for use  
PT in diagnostic assays, detection and vaccines for Norwalk and related

PT viruses.  
XX Claim 14; Page 68-70; 156pp; English.  
XX The Norwalk virus was isolated from stool samples from adult volunteers  
XX infected with safety tested Norwalk virus strain 8FIIa. The coding  
XX sequence is useful for the design of probes for use in diagnostic assays  
XX for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct  
XX PN field.) (Updated on 16-OCT-2003 to standardise OS field)  
XX SQ Sequence 530 AA;

Query Match 99.1%; Score 2793; DB 2; Length 530;  
Best Local Similarity 98.7%; Pred. No. 3.4e-245;  
Matches 523; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MMASKDATSSVDGAGAGQVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINF 60  
DB 1 MMASKDATSSVDGAGAGQVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINF 60

QY 61 VOAQGEFTISPNNTPGGVLFDLSLGLPHLNPFLHLSQMYNGWGNVRVIMLAGNAFTA 120  
DB 61 VOAQGEFTISPNNTPGGVLFDLSLGLPHLNPFLHLSQMYNGWGNVRVIMLAGNAFTA 120

QY 121 GKIIIVSCIPPGFSGSHNLTTIAQATLFPFHVIADVRTLDPIEVLPLEDVNRVLFHNDRNQTM 180  
DB 121 GKIIIVSCIPPGFSGSHNLTTIAQATLFPFHVIADVRTLDPIEVLPLEDVNRVLFHNDRNQTM 180

QY 181 RLVCMLYTLPLRTGGTGGDSFVAGRVMTCPSPDFNLFVPTVEQKTRPFTLPNPLSS 240  
DB 181 RLVCMLYTLPLRTGGTGGDSFVAGRVMTCPSPDFNLFVPTVEQKTRPFTLPNPLSS 240

QY 241 LNSRAPLPISGMISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300  
DB 241 LNSRAPLPISGMISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTVIN 300

QY 301 LTLDGTPFPFPGAPIGFDPDLGGCDWHINMTQFGHSSQTYDVTDPDTFVPHLGSIQ 360  
DB 301 LTLDGTPFPFPGAPIGFDPDLGGCDWHINMTQFGHSSQTYDVTDPDTFVPHLGSIQ 360

QY 361 ANGIGSNYIGVLSWVSPSPHSPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420  
DB 361 ANGIGSNYIGVLSWVSPSPHSPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420

QY 421 FMSKIPGAGYSLPCLLPQEIYISHLASEQAPTVEAALLHYVDPDGTGRTLGEFKAYPDGF 480  
DB 421 FMSKIPGAGYSLPCLLPQEIYISHLASEQAPTVEAALLHYVDPDGTGRTLGEFKAYPDGF 480

QY 481 LTCVPNGASSGQQQLPINGVFVFSWVSRRFYQLKPVGTASSARGRLGLRR 530  
DB 481 LTCVPNGASSGQQQLPINGVFVFSWVSRRFYQLKPVGTASSARGRLGLRR 530

RESULT 3  
ADC72176  
ID ADC72176 standard; protein; 530 AA.  
XX AC ADC72176;  
XX DT 18-DEC-2003 (first entry)  
XX DE Norwalk virus protein 2 amino acid sequence.  
XX KW immune response; non-Norwalk virus agent; immunogen; Norwalk virus;  
XX KW viral pathogen; acute gastroenteritis; virucidal; antiviral vaccine.  
XX OS Norwalk virus.  
XX PN **US6572862-B1**.  
XX PD 03-JUN-2003.  
XX PT 07-JUN-1995; 95US-00486049.

XX 08-NOV-1989; 89US-00433492.  
 PR 27-APR-1990; 90US-00515993.  
 PR 27-AUG-1990; 90US-00573509.  
 PR 06-MAY-1991; 91US-00696454.  
 XX (BAYU ) BAYLOR COLLEGE MEDICINE.  
 PA ~~Estes MK~~ Jiang X, Graham DY;  
 PI WPI; 1994-101125/12.  
 XX N-PSDB; ADC72174.  
 DR Inducing an immune response against non-Norwalk virus agents, comprises  
 XX administering an immunogen recombinantly expressed from a cDNA from  
 PT Norwalk virus.  
 PT Norwalk virus.  
 XX Example 4; SEQ ID NO 3; 45pp; English.  
 XX This invention relates to a novel method of inducing an immune response  
 CC in an individual against Norwalk virus and non-Norwalk virus agents, by  
 CC orally or parenterally administering an immunogen recombinantly expressed  
 CC or synthesised from a cDNA of Norwalk virus given in the specification.  
 CC Norwalk virus is one of the most important viral pathogens, causing acute  
 CC gastroenteritis. The invention may be used for the development of  
 CC compounds with virucidal activity or an antiviral vaccine. The present  
 CC sequence is the amino acid sequence of a protein encoded by the Norwalk  
 CC virus genome of the invention.  
 XX Sequence 530 AA;  
 SQ

Query Match 99.1%; Score 2793; DB 7; Length 530;  
 Best Local Similarity 98.7%; Pred. No. 3.4e-245;  
 Matches 523; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVNPIDPMIINF 60  
 DB 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVNPIDPMIINF 60  
 QY 61 VOAPQGEFTISPNNTPGGVLFDLSLGPLHNPFLHLSQMYNGWGMVRIMLAGNAFTA 120  
 DB 61 VOAPQGEFTISPNNTPGGVLFDLSLGPLHNPFLHLSQMYNGWGMVRIMLAGNAFTA 120  
 QY 121 GKIIIVSCIPPGGSHNLTTIAQATLPHVIADVRTLDPIEVPLEDVRNVLFNHNDNRQQT 180  
 DB 121 GKIIIVSCIPPGGSHNLTTIAQATLPHVIADVRTLDPIEVPLEDVRNVLFNHNDNRQQT 180  
 QY 181 RLVCMLYTLPLRTGGTGDSEFVAVGRVMTCPSPDFNLFVLPPTVEQTRPFTLPNPLSS 240  
 DB 181 RLVCMLYTLPLRTGGTGDSEFVAVGRVMTCPSPDFNLFVLPPTVEQTRPFTLPNPLSS 240  
 QY 241 LNSRAPLPISGMGI SPDNVQSVQFQNGRCTLDGLRGVTTVPVSLSHVAKIRGTSNGTVN 300  
 DB 241 LNSRAPLPISGMGI SPDNVQSVQFQNGRCTLDGLRGVTTVPVSLSHVAKIRGTSNGTVN 300  
 QY 301 LTELDTGTPHPPEGPAPIGFDPGLGGCDWHINMTQFCHSSQTQYDVTDPFVPHLGSIQ 360  
 DB 301 LTELDTGTPHPPEGPAPIGFDPGLGGCDWHINMTQFCHSSQTQYDVTDPFVPHLGSIQ 360  
 QY 361 ANIGSGNVI GVLWSVSPSHSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFEVLVP 420  
 DB 361 ANIGSGNVI GVLWSVSPSHSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFEVLVP 420  
 QY 421 FMSKIPGCGAYSLPCLLPOEYI SHLASEQAPTVGEAALLHYVDPDTRGLTGFCKAYPDGF 480  
 DB 421 FMSKIPGCGAYSLPCLLPOEYI SHLASEQAPTVGEAALLHYVDPDTRGLTGFCKAYPDGF 480  
 QY 481 LTCVPNGASSGQOOLPINGVFWVSWRSFYQLKPVGTASSARGRLGLRR 530  
 DB 481 LTCVPNGASSGQOOLPINGVFWVSWRSFYQLKPVGTASSARGRLGLRR 530

RESULT 4

AAR57091  
 ID AAR57091 standard; protein; 530 AA.  
 XX  
 AC AAR57091;  
 XX  
 DT 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 05-OCT-1994 (first entry)  
 XX  
 DE Small round virus SRSV/KY/89 capsid protein.  
 XX  
 KW pathogen; acute gastroenteritis; food poisoning; seafood contamination;  
 KW diagnostic assay; human calcivirus; small round virus; SRSV; KY89;  
 KW Norwalk virus; capsid protein.  
 XX  
 OS Small round structured virus.  
 XX  
 PN WO9405700-A2.  
 XX  
 PD 17-MAR-1994.  
 XX  
 PF 07-SEP-1993; 93WO-US008447.  
 XX  
 PR 07-SEP-1992; 92US-00941365.  
 XX  
 PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
 XX ~~Estes MK~~ Jiang X, Graham DY;  
 XX WPI; 1994-101125/12.  
 DR N-PSDB; AAO56832.  
 XX  
 PT DNA from Norwalk and related viruses - used for preparing prods. for use  
 PT in diagnostic assays, detection and vaccines for Norwalk and related  
 PT viruses.  
 XX  
 PS Example 7; Fig 13a; 156pp; English.  
 XX  
 CC The known sequence for Norwalk virus was used to obtain the sequence of  
 CC other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool  
 CC from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide  
 CC cDNA sequence includes part of the polymerase region and the capsid  
 CC region of the genome; the deduced amino acid sequences are AAR57092 and  
 CC AAR57091, respectively. Expression of fragments and derivs. of Norwalk-  
 CC related viruses permits development of diagnostic assays to detect  
 CC antibodies, antigens, viral genetic material or antivirals. (Updated on  
 CC 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS  
 CC field.)  
 XX  
 SQ Sequence 530 AA;  
 Query Match 96.2%; Score 2713; DB 2; Length 530;  
 Best Local Similarity 96.8%; Pred. No. 6.6e-238;  
 Matches 513; Conservative 2; Mismatches 15; Indels 0; Gaps 0;

QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVNPIDPMIINF 60  
 DB 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVNPIDPMIINF 60  
 QY 61 VOAPQGEFTISPNNTPGGVLFDLSLGPLHNPFLHLSQMYNGWGMVRIMLAGNAFTA 120  
 DB 61 VOAPQGEFTISPNNTPGGVLFDLSLGPLHNPFLHLSQMYNGWGMVRIMLAGNAFTA 120  
 QY 121 GKIIIVSCIPPGGSHNLTTIAQATLPHVIADVRTLDPIEVPLEDVRNVLFNHNDNRQQT 180  
 DB 121 GKIIIVSCIPPGGSHNLTTIAQATLPHVIADVRTLDPIEVPLEDVRNVLFNHNDNRQQT 180  
 QY 181 RLVCMLYTLPLRTGGTGDSEFVAVGRVMTCPSPDFNLFVLPPTVEQTRPFTLPNPLSS 240  
 DB 181 RLVCMLYTLPLRTGGTGDSEFVAVGRVMTCPSPDFNLFVLPPTVEQTRPFTLPNPLSS 240  
 QY 241 LNSRAPLPISGMGI SPDNVQSVQFQNGRCTLDGLRGVTTVPVSLSHVAKIRGTSNGTVN 300

Db 241 LNSRAPLPISGMISPDNVQSVQFQNGRCTLDGLRGVTPVSLSHVAKIRGTSNGTVIN 300  
QY 301 LTELDTTPHPPEGPAPIGFDPDLGGCDHNNMTQFCHSSQTOYDVTTPDPFVPHLGSIQ 360  
Db 301 LTELDTTPHPPEGPAPIGFDPDLGGCDHNNMTQFCHSSQTOYDVTTPDPFVPHLGSIQ 360  
QY 361 ANGISGNYIGVLSWVSPSHPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420  
Db 361 ANGISGNYIGVLSWVSPSHPSGQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420  
QY 421 FMSKIPGCGAYSLPCLLPQYISHLASEQAPTVGEAALLHYVDPDTGRTLGFEKAYPDGF 480  
Db 421 FMSKIPGCGDSLPCLLPQYISHLASEQAPTVGEGLLHYVDPDTDRNLGEFKAYPDGF 480  
QY 481 LTCVNGASSGQQLPPIGVFVSVWVSRFYQLKPVGTASSARGRLGLRR 530  
Db 481 LTCVNGASSGQQLPPIGVFVSVWVSRFYQLKPVGTASTARGRLGLRR 530  
  
RESULT 5  
AAB49703  
ID AAB49703 standard; protein; 544 AA.  
XX  
AC AAB49703;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 4.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX  
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
XX  
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX  
DR WPI; 2001-080848/09.  
DR N-PSDB; AAF29144.  
XX  
PT Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX  
PS Claim 1; Page 47-49; 84pp; Japanese.  
XX  
CC This invention relates to a kit for the detection and typing of small  
CC round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX  
SQ Sequence 544 AA;  
  
Query Match 74.2%; Score 2093; DB 4; Length 544;  
Best Local Similarity 72.7%; Pred. No. 2.1e-181;  
Matches 396; Conservative 53; Mismatches 80; Indels 16; Gaps 8;  
  
QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLANDPVAGSSSTAVATAGVNPIDPWLIINN 60  
Db 1 MMASKDATSPADGATGAGQLVPEVNTADPIPDPVAGSSSTALATAGVNLIDPWLIINN 60

QY 61 VQAPQGBFTTISPNNTPGGVFLFDLSGLPHLAPFLHLHLSQMYNGWGNVRIMLAGNAFTA 120  
Db 61 VQAPQGBFTTISPNNTPGGVFLFDLSGLPHLAPFLHLHLSQMYNGWGNVRIMLAGNAFTA 120  
QY 121 GKIVSCIPPGFGGHNUTIAQATLFPHVIAADVRLTDIEVPLEDVRVNLPHNNRNQOTM 180  
Db 121 GKVIICCPVPGFQSRILSIQAATLFPHVIAADVRLDPEVEPLEDVRVNLPHNNRNQOTM 179  
QY 181 RLVCMLVTPRLTGGGTG--DSFVVAGVMTCPSPDENFLFLVPPVTEQKTRPFTLPNLPL 238  
Db 180 RLLCMLVTPRLTGGGTGDSFVVAGVMTCPSPDENFLFLVPPVTEQKTRPFTLPNLPL 239  
QY 239 SSLNSRAPLPISGMISPDNVQSVQFQNGRCTLDGLRGVTPVSLSHVAKIRG--TSNG 296  
Db 240 KYLSNSRIPNPIEGMSLSPDQTNVQFQNGRCTIDGQPLGTTVPVSVSLCKFRGRITSGQ 299  
QY 297 TVINLTLDGTPHPPEGPAPIGFDPDLGGCDHNNMTQFCHSSQTOYDVTTPDPFVPHL 351  
Db 300 RVLNLTLDGSPFMAFAAPAPAGPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQQ 359  
QY 352 FVPHLGSI--QANGIGSGNYIGVLSWVSPSHPSGQVDLWKIPNYGSSITEATHLAPSV 409  
Db 360 FVPHLSSITLDENVSSGGDIYIGTQWTSPPSDSGGANTFNFKIPDYGSSLAESQLAPAV 419  
QY 410 YPPGFGEVLVFFMSKIPGP---GAYSL-PCLLPQYISHLASEQAPTVGEAALLHYVDPD 465  
Db 420 YPPGFNEVIVFMASIPGPNQSGPNLVPCLLPQYITHFISEQAPIQGEAALLHYVDPD 479  
QY 466 TGRITLGEFKAYPDGFLTCVNGASSGQQLPPIGVFVSVWVSRFYQLKPVGTASSARGR 525  
Db 480 TNRNLGEFKLPGGYLTCVPSNSTGPPQQLPLDGVFVSVWVSRFYQLKPVGTAGPARGR 539  
QY 526 LGLRR 530  
Db 540 LGVRR 544  
  
RESULT 6  
AAB49702  
ID AAB49702 standard; protein; 546 AA.  
XX  
AC AAB49702;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 3.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX  
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
XX  
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX  
DR WPI; 2001-080848/09.  
DR N-PSDB; AAF29143.  
XX  
PT Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX  
PS Claim 1; Page 45-47; 84pp; Japanese.  
XX



CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX  
 XX Sequence 546 AA;

Query Match 72.0%; Score 2029.5; DB 4; Length 546;  
 Best Local Similarity 69.4%; Pred. No. 1.3e-175;  
 Matches 379; Conservative 68; Mismatches 82; Indels 17; Gaps 6;

QY 1 MMWASKDATSSVDGASGAGQLVPEYNASDPLAMPVAGSSTAVATAGVQNPIDPIMNNF 60  
 DB 1 MMWASKDAQSDAGSAGAGQLVPEYNTADPLMPVAGPTTAVATAGVQNMIDPIMVNF 60  
 QY 61 VQAPQGEFTISPNNTPGGVLDLSLGPLHNPFLHLSQMYNGWGMVRVIRIMLAGNAFTA 120  
 DB 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLNPFLHLSQMYNGWGMVRVIRILLAGNAFSA 120  
 QY 121 GKIIIVCIPPGSGSHNLTTAQATLPHVIAVRTLDPIEVPLEDVNRVLFHNNDRNQOTM 180  
 DB 121 GKIIIVCCVPPGFTSSSLTTAQATLPHVIAVRTLEPIEMPLEDVNRVLYHTND-NQPTM 179  
 QY 181 RLVCMLYTLPLRTGGGTG--DSFVAVAGRMTCSPDFNLFVLPVPTVEQKTRPFTLPNPL 238  
 DB 180 RLVCMLYTLPLRTGGSGNSDSFVAVAGRLTAPSSDFSLFVLPPTIEQKTRPFTVPNPL 239  
 QY 239 SLSNSRABLPISGMGISPDNVQSFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSN--G 296  
 DB 240 QTLNSRFPESLIQGMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINOGA 299  
 QY 297 TWINLTGDTPEHPPEGPAPIGFDPDLGGCDWHINNTQFGHSSQT-----QYDVTPTPT 351  
 DB 300 RTUNLTFEVDGKFMAFDSFAPVGFDFGKCDWHMRISKTPNNTSSGDPMSRVSVQTNVQG 359  
 QY 352 FVPHLGSIQANGI--GSGNYIGVLSWVSPSPHSPGQVDLWKIPNYGSSITEATHLAPS 408  
 DB 360 FVPHLGSIQDFVFNHTPDYIGTIEWISQPTPTGTDINLWEIPDYGSSLSQAANLAPP 419  
 QY 409 VTPPGGEVLVFMFKIPGFGAYS----LPCLLPQYIISHLASEQAPTVGEAALLHYVDP 464  
 DB 420 VPPPGGEALVYFVSAPFGPNRNSAPNDVPCLLPQYIITHFVSEQAPTMGDAALLHYVDP 479  
 QY 465 DTGRTLGEFKAYPDGFLTCVPGNASSGPOOLPINGVVFVSVWSRYQLKPVGTASSARG 524  
 DB 480 DTRNLGEBFKLYPGGYLTCVPGNGVAGAPOLPLNGVFLFVSVWSRYQLKPVGTASTARS 539  
 QY 525 RLGLRR 530  
 DB 540 RLGVRR 545

## RESULT 7

AAB49700  
 ID AAB49700 standard; protein; 545 AA.

XX AAB49700;

AC AAB49700;

XX 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 1.

XX Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.

XX WO200079280-A1.

XX 28-DEC-2000.

PF 22-JUN-2000; 2000WO-JP004095.  
 XX  
 PR 22-JUN-1999; 99JP-00175928.  
 XX  
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 PA (DENK-) DENKA SEIKEN KK.  
 XX  
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 DR WPI; 2001-080848/09.  
 DR N-PSDB; AAF29141.  
 XX  
 PT Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.  
 XX  
 PS Claim 1; Page 40-42; 84pp; Japanese.  
 XX  
 CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX  
 SQ Sequence 545 AA;

Query Match 68.7%; Score 1938; DB 4; Length 545;  
 Best Local Similarity 67.0%; Pred. No. 2.7e-167;  
 Matches 366; Conservative 63; Mismatches 99; Indels 18; Gaps 7;

QY 1 MMWASKDATSSVDGASGAGQLVPEYNASDPLAMPVAGSSTAVATAGVQNPIDPIMNNF 60  
 DB 1 MMWASKDAPTNDGTSGAGQLVPEANTAEPISEMPVAGAAATAATAGVQNMIDPIMNNY 60  
 QY 61 VQAPQGEFTISPNNTPGGVLDLSLGPLHNPFLHLSQMYNGWGMVRVIRIMLAGNAFTA 120  
 DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLHLSQMYNGWGMVKVYLLAGNAFTA 120  
 QY 121 GKIIIVCIPPGSGSHNLTTAQATLPHVIAVRTLDPIEVPLEDVNRVLFHNNDRNQOTM 180  
 DB 121 GKIIIVCIPPGGPAANISIAQATMPHVIADVRLVLEIEVPLEDVNRVLFHNNDRNQOTM 179  
 QY 181 RLVCMLYTLPLRTGGGTG--DSFVAVAGRMTCSPDFNLFVLPVPTVEQKTRPFTLPNPL 238  
 DB 180 RLVCMLYTLPLRAGSSSGTDPFVIAGRLVTCPSDFSLFVLPVPPVNEQKTRPFTLPNPL 239  
 QY 239 SLSNSRABLPISGMGISPDNVQSFQNGRCTLDGRLVGTTPVSLSHVAKIRGT-----S 294  
 DB 240 NTLNSRVPESLTKSMVSRDRHQMVQFQNGRVRTLDGQLGTTPTTSASQLCKIRGVSFHAN 299  
 QY 295 NGTVINLTGDTPEHPPEGPAPIGFDPDLGGCDWHIN--MTQFGHSSQTQYDVTPTPT 351  
 DB 300 GNGYNLTGDSGYHAFESPAPIGFDPDLGBCDWHMEASPTTQFNTGDKVQINVKQBSA 359  
 QY 352 FVPHLGSIQANGIG---SGNYIGVLSWVSPSPHSPGQVDLWKIPNYGSSITEATHLAP 407  
 DB 360 FAPHLGTIQADGLSDVSVNTNMIKLGWVSPVSDGHRGDVDPWVIPRYGSTLITAAQLAP 419  
 QY 408 SVYPPGGEVLVFMFKIP---GPGAYSILPCLLPQYIISHLASEQAPTVGEAALLHYVDP 464  
 DB 420 PIYPPGGEALVYFMSDFPIAHGTNGLSVPCITPQEFVTHFVNEQAPTRGEAALLHYLDP 479  
 QY 465 DTGRTLGEFKAYPDGFLTCVPGNASSGPOOLPINGVVFVSVWSRYQLKPVGTASSARG 524  
 DB 480 DTRNLGEBFKLYPEGFMTCVPNSGCTGPTUPINGVVFVSVWSRYQLKPVGTAGPA-C 538  
 QY 525 RLGLRR 530  
 DB 539 RLGIIR 544

RESULT 8  
AAB49706  
ID AAB49706 standard; protein; 540 AA.  
XX  
AC AAB49706;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 7.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX  
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX (DENK-) DENKA SEIKEN KK.  
XX  
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX  
DR WPI; 2001-080848/09.  
XX N-PSDB; AAF29147.  
XX  
PT Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX  
XX Claim 1; Page 54-57; 84pp; Japanese.  
XX  
XX This invention relates to a kit for the detection and typing of small  
XX round structured virus (SRSV) strains. The kit contains antibodies  
XX directed against peptides represented in sequences AAB49700 - AAB49710,  
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
XX used for detecting and typing strains of SRSV in order to prevent the  
XX spread of infection and to examine the epidemiology of outbreaks  
XX  
XX Sequence 540 AA;  
Query Match 43.7%; Score 1232; DB 4; Length 540;  
Best Local Similarity 47.3%; Pred. No. 5.6e-103;  
Matches 264; Conservative 81; Mismatches 157; Indels 56; Gaps 12;  
QY 1 MMMAKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSSAVATAGQVNPIDPWIIINF 60  
DB 1 MKMASNDATPSNDGAAG---LVPESN-NEAMALEFVVGASLAAPVTGQTIIIDPWIRNF 56  
QY 61 VQAPQGEFTISPNNTPGGVLDLSLGPLNPFLLHLSQMYNGWGNMRVRIMLAGNAFTA 120  
DB 57 VQAPNGEFTVSPRNSFGBILVNLEGLPBNPFLAHLARMYNGYAGMEVQVMLAGNAFTA 116  
QY 121 GKIIIVSCIPPGFGSHNLTAQATLPPHYIADVRTLDPIEVLPEVDVRNVLFHNDNRNOQT 180  
DB 117 GKIIIFAAVPPFPVENLSPSQITMFPFHVIIIDVRLTEPVLLEPMPDVRSTLFEHNOKEPKM 176  
QY 181 RLVCMLYTLRTGGTGSFVAGRWMTCPSPDENFLFVPTVEOKTRPTLNLPLSS 240  
DB 177 RLVMALYTLPLRNSGSGDDVFTVSCRILTRSPGEPDFTYLVPTVESKTKPTFLPVLTIGE 236  
QY 241 LNSNRAPLPISGMGISPNQVQFQNGRCTLGDLGVLGTTPVSLSHVAKIRGTNGTV-- 298  
DB 237 LNSRFPPLSIDEMVTSPIESIVVQPNGRVTLGELLGTTLQACNICIRGKVTGQVPS 296  
QY 299 -----INLTDLGTPFHPFEG-PAPIGFPDLGGCDWHINMTQFGHSSQTQ----- 342  
DB 297 EQHMWNLEITNLNGTQFDPTDDVPAPLGPVPDAG-----EVFGVLSQRNRRGNSPANR 349

QY 343 -YD--VDTPDTFVPHLSIQANGISGNYIGV-----LSWVSPPHSPSGSQVD 388  
DB 350 AHDAVATYSDKYTKPLGLVQ---IGTWNTDVENQTKFTPIGLNEVA-----NGHRPE 401  
QY 389 LWKIPNYGSSITRATHLAPSVYPPGFGVLFVFMFSKIPGPGAYSLLP---CLLPQEVLSHL 445  
DB 402 QWTLPRYSGALTINMNLAPAVAPLFPGERLLFPFRSYVPLKGGFGNPAIDCSVPQEWQHF 461  
QY 446 ASEQAPTVGEAALLHYVDPTGRTIGCFKAYPDGFLTCVPPNGASSGPQOLPINGVVFVS 505  
DB 462 YQESAPSLGDVALVRYVNPDTGRVLFKHLKGGFLTV--SSTSTGTVVVPANGYKFD 519  
QY 506 WWSRFYOLKPVGTASSAR 523  
DB 520 WVNQFYSLAPMGTNGRR 537  
RESULT 9  
AAB49705  
ID AAB49705 standard; protein; 548 AA.  
XX  
AC AAB49705;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 6.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX (DENK-) DENKA SEIKEN KK.  
XX  
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX  
DR WPI; 2001-080848/09.  
XX N-PSDB; AAF29146.  
XX  
PT Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX  
XX Claim 1; Page 52-54; 84pp; Japanese.  
XX  
XX This invention relates to a kit for the detection and typing of small  
XX round structured virus (SRSV) strains. The kit contains antibodies  
XX directed against peptides represented in sequences AAB49700 - AAB49710,  
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
XX used for detecting and typing strains of SRSV in order to prevent the  
XX spread of infection and to examine the epidemiology of outbreaks  
XX  
XX Sequence 548 AA;  
Query Match 42.9%; Score 1209; DB 4; Length 548;  
Best Local Similarity 45.4%; Pred. No. 7.1e-101;  
Matches 260; Conservative 72; Mismatches 163; Indels 78; Gaps 11;  
QY 1 MMMAKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSSAVATAGQVNPIDPWIIINF 60  
DB 1 MKMASNDATPSNDGAAG---LVPEIN-NEAMALDPVAGAAIAAPLTGQONIIDPWIMNF 56  
QY 61 VQAPQGEFTISPNNTPGGVLDLSLGPLNPFLLHLSQMYNGWGNMRVRIMLAGNAFTA 120

Db 57 VQAPGGEFTVSRNPSGEGVLLNLELCPENIPFLAHLMRYNGYAGGFEVQVVLGNAFTA 116  
Qy 121 GKIIIVSCIPPGGSHNLTAQATLPHVIAVDTLDPIEVPLEDVNRNVLPHNNDRNQOQM 180  
Db 117 GKIIIPAAIPNPIDNLNSAAQITMCPHVIVDRQLEFVNLPMFVDRNFFHYNQGSDSL 176  
Qy 181 RLVCMLYTLPLRTGGTGDSEFVAGRVMTCPSPDFNPLFLVPTVQKTRPFTLPLNLPLSS 240  
Db 177 RLIAMLYTLPLRANSGDDVFTVSCRVLTRPSDFSNFLVPTVBSKTKPFTLPLITISE 236  
Qy 241 LNSRAPLPISGMGLSPDNVQSFQNGRCTLDGRLVGTTPVSLSHVAKIRG-----TSN 295  
Db 237 MNSRFPVPIESLHTSPTENIVVQCGRVTLIDGELMGTTQLPLSQICAFRGVLRSTRS 296  
Qy 296 GT-----VINLTLDGTPPHPEG-PAPIGFPLDGLGCDWHINMTQFGHSS 339  
Db 297 ASDQADATPRLFNYYVHVLQNLNGTPYDPAEDIPGLGTDFRG-----KVFQVAS 349  
Qy 340 QTOYD-----VDTTPTDTPVPHLGSIOAN-----GIGSGNYIGVL 373  
Db 350 QRLNLDSTTRAHAKVDTTAGRFTPKLGSLEISTDSDDFQNGQTKFTPVGIGVDN----- 404  
Qy 374 SVNSPSPSGSQVDLWKIPNYGSSITEATHLAPSVYPPGCEVLVFPFMSKIPGPCAYS- 432  
Db 405 -----EAFQWMSLPDYSGOFTHMNLAPAVAPNFPGEQLLFFRSQLPSSGGRSN 454  
Qy 433 --LPCLLPOEYIASHLASEQAPTVGEAALHYVDPTDGTGLGFKAYPDGFLTCVPGNASS 490  
Db 455 GVLDCVLPOEWQHFQIESAPAQTOVALVRYVNPDTGKVLFEAKLHKLGFMTIANGDS- 513  
Qy 491 GQQQLPINGVVFVSVSVSRFYOLKPVGTASSAR 523  
Db 514 -FITVPPNGYFRFESWVNPFTLAPMGTGNRR 545

RESULT 10  
ID AAB49707  
XX AAB49707 standard; protein; 535 AA.  
AC AAB49707;  
XX 04-APR-2001 (first entry)  
DT Small round structured virus protein SEQ ID 8.  
DE Small round structured virus; SRSV; food poisoning.  
KW Small round structured virus.  
OS Small round structured virus.  
XX WO200079280-A1.  
XX 28-DEC-2000.  
XX 22-JUN-2000; 2000WO-JP004095.  
XX 22-JUN-1999; 99JP-00175928.  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX (DENK-) DENKA SEIKEN KK.  
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX WPI: 2001-080848/09.  
XX N-PSDB; AAF29148.  
XX Kit for the detection and typing of small round-structured virus (SRSV)  
XX PT strains for investigation of food poisoning outbreaks, contains  
XX antibodies.  
XX Claim 1; Page 57-59; 84pp; Japanese.  
XX This invention relates to a kit for the detection and typing of small

CC round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX Sequence 535 AA;  
SQ Query Match 42.6%; Score 1200.5; DB 4; Length 535;  
Best Local Similarity 46.4%; Pred. No. 4.1e-100;  
Matches 258; Conservative 76; Mismatches 165; Indels 57; Gaps 12;  
Qy 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVNPIDPMIINF 60  
Db 1 MMASNDAAAPSNDGAAG---LVPEAN-NETWALEPVAGASIAAPLTGQNNIIDPMIRLNF 56  
Qy 61 VQAPGEBETISPNNTPGGVLPDLISGPHLNPPLHLSQMYNGVGMVRVIMLAGNAFTA 120  
Db 57 VQAPNGEFTVSRNPSGEGVLLNLELCPENIPFLAHLSRMYNGYAGGFEVQVVLGNAFTA 116  
Qy 121 GKIIIVSCIPPGGSHNLTAQATLPHVIAVDTLDPIEVPLEDVNRNVLPHNNDRNQOQM 180  
Db 117 GKLVFAAAPPHPPLENISPGQITMFPVHVIDVRLTLEPVLLPLPDVRRNFFHYNQOQPRM 176  
Qy 181 RLVCMLYTLPLRTGGTGDSEFVAGRVMTCPSPDFNPLFLVPTVQKTRPFTLPLNLPLSS 240  
Db 177 RLVCMLYTLPLRANSGDDVFTVSCRVLTRPSDFSNFLVPTVBSKTKPFTLPLITIGE 236  
Qy 241 LNSRAPLPISGMGLSPDNVQSFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNGTV-- 298  
Db 237 LTNSRFPVPIDELYTSPNESLVQPNQGRCALDGELOQTQLLPTAICSFGRGRINQKVS 296  
Qy 299 -----INLTLDGTPPHPE-EGPAPIGFPLDGLGCDWHINMTQFGHSSQTYD----- 344  
Db 297 ENHVMVMQVTININGTPFDPTGDVPAPLGTDFSG-----KLFGLVSLQRDHDAACRSHD 349  
Qy 345 --VDTTPTDTPVPHLGSIOANGIGSGNYIGVLSWSVSPSH-----PSG-----SQVDLW 390  
Db 350 AVIATNSAKFTPKLGAIQ-----IG--TWBDDVHINQPTKFTPVGLPENEGFNQW 398  
Qy 391 KIPNYGSSITEATHLAPSVYPPGCEVLVFPFMSKIPGPCAYSPL---CLLPOEYIASHLAS 447  
Db 399 TLPNYSGALTLNMLGAPPVAPTFFPGEQLLFFRSHIPKGGVADPVIDCLLPOEYIASHLQ 458  
Qy 448 EOAPTVGEAALHYVDPTDGTGLGFKAYPDGFLTCVPGNASSGPPQLPINGVVFVSVW 507  
Db 459 ESAPSQSDVALIRFTNPDGTGRVLFEAKLHRSYIIVANTG--SRFIVVPANGYFRFDTWV 516  
Qy 508 SRFYOLKEVGTASSAR 523  
Db 517 NOFYSLAPMGTGNRR 532  
RESULT 11  
ID AAU91272  
XX AAU91272 standard; protein; 548 AA.  
AC AAU91272;  
XX 18-JUN-2002 (first entry)  
DT Norwalk virus associated polynucleotide #1.  
DE Norwalk virus associated polynucleotide #1.  
KW Norwalk virus; monoclonal antibody; geno group I; geno group II;  
KW immunological detection; food; viral infection.  
OS Norwalk virus.  
XX JP2002020399-A.  
XX 23-JAN-2002.  
XX

PF 10-JUL-2000; 2000JP-00208151.  
 FR 10-JUL-2000; 2000JP-00208151.  
 XX (OSAP ) OSAKA PREFECTURE.  
 PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.  
 PA (IATR ) IATRON LAB INC.  
 XX WPI; 2002-287412/33.  
 XX A monoclonal antibody useful in the immunological detection and diagnosis  
 PT of Norwalk virus infection.  
 XX Disclosure; Page 12-13; 24pp; Japanese.  
 PS The invention describes a monoclonal antibody recognising Norwalk virus,  
 CC a capsid protein of Norwalk virus, or a common antigen epitope on the  
 CC capsid protein molecule of geno group I and geno group II. The antibody  
 CC is useful for immunological detection and quantitative analysis of  
 CC Norwalk virus in foods and the serum of infected patients. This sequence  
 CC represents a Norwalk virus associated protein described in the invention  
 XX  
 XX Sequence 548 AA;  
 SQ Query Match 42.1%; Score 1188; DB 5; Length 548;  
 Best Local Similarity 45.3%; Pred. No. 5.8e-99;  
 Matches 258; Conservative 70; Mismatches 164; Indels 78; Gaps 11;  
 QY 1 MMWASKDATSSVDGASGAGQGVLPVNASDPLAMDVPVAGSSTAVATAGQVNPIDPWIINN 60  
 DB 1 MKWASRAAFNSDGAAG---LVPEIN-NEAALDPVAGAAIAAPLTGQQNIIDPWIINN 56  
 QY 61 VQAPQGEFTTSPNNTPGGVLPDLSLGHLPNFPFLHLSOMYNGWGNVRVIMLAGNAFTA 120  
 DB 57 VQAPGGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGGFEVQVVLGNAFTA 116  
 QY 121 GKIIIVSCIPGFGSHNLTIAQATLPHVIAVDRITLPIEVPLEDVRNVLPHNDRNQTM 180  
 DB 117 GKIIIFAAIPNFPIDNLSAAQITMCPHVIIVDRQLEPNLPMVVRNFPVHYNQGSRL 176  
 QY 181 RLVCMLYTLRTGGTGDGDFVAGRVMTCPSPDFNLFVLPVPTVEQKTRPTLPLNPLSS 240  
 DB 177 RLIAMLYTLPLRANNSGDDVFTVSCVLTTRSPDFSFNLFVPTVESKTKPTLPLTISE 236  
 QY 241 LNSRAPLPIGSGMISPNVQSFQNGRCITLDELGLVGTTPVSLSHVAKIRG-----TSN 295  
 DB 237 MNSRFPVPIESLHTSPTENIIVQCNQGRVTLDELGLVGTTPVSLSHVAKIRG-----TSN 296  
 QY 296 GT-----VINLTLDGTPFPFEG-PAPIGPDLGGCDWHNNMTQFGHSS 339  
 DB 297 ASDQADIATPLFNYYWHVQJONLNGTTPYDPAEDIPGLGTPDRG-----KVFGVAS 349  
 QY 340 QTQYD-----VDTTPTDFVPHLGSIQAN-----GIGSGNYIGVL 373  
 DB 350 QENPSTTTRAEAKVDYTAGRTPKLGSLEISTESSDPDQNTQTRTPVGVGDV-----404  
 QY 374 SWVSPSPHSGQVDLWKIPNGSSITEATHLAPSVYPPGFEVLVFMKIPGPGAYS- 432  
 DB 405 -----EADFOQSLPDYSGQFTNNMLAPAVAFNFGEEQLLFRSLPSSGGRSN 454  
 QY 433 --LPCLLPQYIYSLHSAEQAPTVEAALLHYVDPDTRTLGFEKAYPDGFLTCVPNGASS 490  
 DB 455 GVLDCLVPOEWQHFYQESAPAQTOVALVRYVNPDTGRVLPEAKLHLKLGFTIAKNGDS- 513  
 QY 491 GPQQLPINGVVFVMSVSRFYOLKPVGTAS 520  
 DB 514 -FITVPPNGYFRFESWVNPFTLAPMGTCN 542  
 RESULT 12  
 ID AAB49708  
 ID AAB49708 standard; protein; 542 AA.  
 XX

AAB49708;  
 04-APR-2001 (first entry)  
 Small round structured virus protein SEQ ID 9.  
 Small round structured virus; SRSV; food poisoning.  
 Small round structured virus.  
 WO200079280-A1.  
 28-DEC-2000.  
 22-JUN-2000; 2000WO-JP004095.  
 22-JUN-1999; 99JP-00175928.  
 (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 (DENK-) DENKA SEIKEN KK.  
 Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 WPI; 2001-080848/09.  
 N-PSDB; AAF29149.  
 Kit for the detection and typing of small round-structured virus (SRSV)  
 strains for investigation of food poisoning outbreaks, contains  
 antibodies.  
 Claim 1; Page 59-61; 84pp; Japanese.  
 This invention relates to a kit for the detection and typing of small  
 round structured virus (SRSV) strains. The kit contains antibodies  
 directed against peptides represented in sequences AAB49700 - AAB49710,  
 which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 used for detecting and typing strains of SRSV in order to prevent the  
 spread of infection and to examine the epidemiology of outbreaks

Query Match 41.9%; Score 1181; DB 4; Length 542;  
 Best Local Similarity 46.7%; Pred. No. 2.5e-98;  
 Matches 259; Conservative 82; Mismatches 166; Indels 48; Gaps 14;  
 QY 1 MMWASKDATSSVDGASGAGQGVLPV-VNASDPLAMDVPVAGSSTAVATAGQVNPIDPWIINN 59  
 DB 1 MKWASNDAAPSNDGAA---SLVPEGINETMPL--EPVAGASIAAPVAGQTNIIDPWIRTN 55  
 QY 60 FVQAPQGEFTTSPNNTPGGVLPDLSLGHLPNFPFLHLSOMYNGWGNVRVIMLAGNAFTA 119  
 DB 56 FVQAPNGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGGFEVQVVLGNAFTA 115  
 QY 120 AGKIIIVSCIPGFGSHNLTIAQATLPHVIAVDRITLPIEVPLEDVRNVLPHNDRNQOT 179  
 DB 116 AGKILFAAIPNFPIDNLSAAQITMCPHVIIVDRITLPIEVPLEDVRNVLPHNDRNQOT 175  
 QY 180 MRLVCMYTLRTGGTGDGDFVAGRVMTCPSPDFNLFVLPVPTVEQKTRPTLPLNPLS 239  
 DB 176 MRLVAMLYTLPLRANNSGDDVFTVSCVLTTRSPDFSFNLFVPTVESKTKPTLPLTISE 235  
 QY 240 SLNSRAPLPIGSGMISPNVQSFQNGRCITLDELGLVGTTPVSLSHVAKIRG---TSGN 296  
 DB 236 ELTNSRFPPIEQLYTAPNETVNVQCNQGRVTLDELGLVGTTPVSLSHVAKIRG---TSGN 295  
 QY 297 -----TVINLTLDGTPFPFEG-PAPIGPDLGGCDW-----HNNMTQFGHSSQTQYD 344  
 DB 296 DNWDQNLQLTYENGASYDTPDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGY 355  
 QY 345 VDTTPTDFVPHLGSIQANIGSGNYIGVLVSWSPSH-----PSSGQVD-----LWKIPN 394  
 DB 356 ISTSGKFTPKIGSI---GLHS-----ITEHVHNPQOSRFTPVGVAVDENTFPQWVLPH 407

QY 395 YGSSITEATHLAPSVYPPGFGVLFVFMKIP-----GPGAYSILPCLLPQBYISHLASE 448  
DB 408 YAGSLALNTLAPAVAPTTPGSQLFFRSRVCVQGLQCDAP-IDCLLPQEWVNFYQE 466  
QY 449 QAPTVEGALHYVDPDGTGRTLGEFKAYPDGFLTCVPGNCASSGPOOLPINGVVFVSVWS 508  
DB 467 AAPSQADVALIRYNPDGTGRTLFEAKLHRSFGTIVSHGTAY--PLVVPNGHFRFDSWN 524  
QY 509 REYQLKPVGTASSAR 523  
DB 525 QFYSLAPMGTGNRR 539  
  
RESULT 13  
ID AAB49704 standard; protein; 539 AA.  
XX AAB49704;  
AC AAB49704;  
DT 04-APR-2001 (first entry)  
XX Small round structured virus protein SEQ ID 5.  
DE Small round structured virus; SRSV; food poisoning.  
KW Small round structured virus.  
OS Small round structured virus.  
XX WO200079280-A1.  
PN 28-DEC-2000.  
PF 22-JUN-2000; 2000WO-JP004095.  
PR 22-JUN-1999; 99JP-00175928.  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
DR WPI; 2001-080848/09.  
DR N-PSDB; AAF29145.  
XX Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX Claim 1; Page 50-52; 84pp; Japanese.  
XX This invention relates to a kit for the detection and typing of small  
CC round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX  
SQ Sequence 539 AA;  
  
Query Match 40.7%; Score 1148.5; DB 4; Length 539;  
Best Local Similarity 44.8%; Pred. No. 2.3e-95;  
Matches 247; Conservative 85; Mismatches 182; Indels 37; Gaps 14;  
  
QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVAGSSTAVATAGQVNDPDIINNF 60  
DB 1 MMASNDANPS-DGST--ANLVPEYN-NEVMALEFVVGAAIAAPVAGQNVDPWIRNMF 56  
  
QY 61 VOAPQGEFTISPNTPGGVLPDLSPHLPNPLHLSQYNGWGNVRIMLAGNAFTA 120  
DB 57 VOAPGGEFTVSRNAPGILMSAPLCPDLNPNVLSHARMYNGYAGGFVQVILAGNAFTA 116  
  
QY 121 GKIIIVSCIPPGFGSHNLTAQATLPPHVIADVRTLDPIEVPLEVDRNVLFHNDRNQOTM 180

DB 117 GKIIFAAAPPFPTEGLSPSQVTMPPHIIVDRQLPVLILPLPDRNNFYHYNQNDSTI 176  
QY 181 RLVCMLYTPRLTGGTGDGDFVAGRVMTCPSPDFNLFVLPVPTVQKTRPFTLPNLPSS 240  
DB 177 KLIAMLYTPLRANAGDDVFTVSCRVLTRPSDFDFILVPTVESRTKPTVPILTVEE 236  
QY 241 LSNRAPILPISGMGISPDNVQSQNGRCCTLDGRLVGTTPVSLSHVAKIRGT-----S 294  
DB 237 MSNRFPIPLEKLYTGPSSAFVQVQNGRCCTTDGVLGTTQLSAVNICTFRGDVTIAGS 296  
QY 295 NGTVNLTELDGTPHPPEG-PAPIGFPD-LGGCDWHINMT--QFGHSSQQTQYDVTTPD 350  
DB 297 HDYTNLASQWNSNYDPTTEIPAPLGTDFVKIGQMLTQTTREDGSTRAHKATVSTGSSV 356  
QY 351 TFVPHLGSIQ-----ANGIGSGNY-----IGVLSWVSPSPSGSQVDLWKIPNYGSSIT 400  
DB 357 HFTPKLGSVQYTTDNDNFQGTNTKFTVPVGIQ--DGNHQNPEQ--QWVLPNYSGRGT 412  
QY 401 BATHLAPSVYPPGFGVLFVFMKIPGPGAY---SLPCLLPQBYISHLASEOAPTVEAA 457  
DB 413 HNVHLAPAVAPTTPGSQLFFRSRVCVQGLQCDAP-IDCLLPQEWVNFYQE 472  
QY 458 LLHYVDPDGTGRTLGEFKAYPDGFLTCVPGNCASSGPOOL--PINGVVFVSVWSRVOLKP 515  
DB 473 LLRFVNPDTGRVLFECKLHKSQYVTV----AHTGPHDLVIPPNGYFRFDSWVNOQVYTLAP 528  
QY 516 VGTASSARGRL 526  
DB 529 MONGNRRAL 539  
  
RESULT 14  
ID AAB49709 standard; protein; 550 AA.  
XX AAB49709;  
AC AAB49709;  
DT 04-APR-2001 (first entry)  
XX Small round structured virus protein SEQ ID 10.  
DE Small round structured virus; SRSV; food poisoning.  
KW Small round structured virus.  
OS Small round structured virus.  
XX WO200079280-A1.  
PN 28-DEC-2000.  
PF 22-JUN-2000; 2000WO-JP004095.  
PR 22-JUN-1999; 99JP-00175928.  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
DR WPI; 2001-080848/09.  
DR N-PSDB; AAF29150.  
XX Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX Claim 1; Page 62-64; 84pp; Japanese.  
XX This invention relates to a kit for the detection and typing of small  
CC round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX  
SQ Sequence 539 AA;  
  
Query Match 40.7%; Score 1148.5; DB 4; Length 539;  
Best Local Similarity 44.8%; Pred. No. 2.3e-95;  
Matches 247; Conservative 85; Mismatches 182; Indels 37; Gaps 14;  
  
QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVAGSSTAVATAGQVNDPDIINNF 60  
DB 1 MMASNDANPS-DGST--ANLVPEYN-NEVMALEFVVGAAIAAPVAGQNVDPWIRNMF 56  
  
QY 61 VOAPQGEFTISPNTPGGVLPDLSPHLPNPLHLSQYNGWGNVRIMLAGNAFTA 120  
DB 57 VOAPGGEFTVSRNAPGILMSAPLCPDLNPNVLSHARMYNGYAGGFVQVILAGNAFTA 116  
  
QY 121 GKIIIVSCIPPGFGSHNLTAQATLPPHVIADVRTLDPIEVPLEVDRNVLFHNDRNQOTM 180

CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX  
 SQ Sequence 550 AA;

Query Match 40.1%; Score 1131; DB 4; Length 550;  
 Best Local Similarity 43.2%; Pred. No. 9.2e-94; Indels 78; Gaps 12;  
 Matches 248; Conservative 77; Mismatches 171; Indels 78; Gaps 12;

QY 1 MMASDKATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIINNF 60  
 DB 1 MKMASNDAAAPSSDGAAG--LVPEAN-DEVWALEPVVGASIAAPVVGQNIIDPWIRENF 56

QY 61 VQAPQGEFTISPNNTPGGVLFDSLGPHNLPFLHLSQMYNGWGNMVRIMLAGNAFTA 120  
 DB 57 VQAPQGEFTVSPNPGEMLLNLELGPENLPYLSHLSRMYNGAGCMQVQVVLGNAFTA 116

QY 121 GKIIIVCIPGFGSHNLTIAQATLPHVIADVRTLDPIEVLDPEDVRNVLFNHNDROOQM 180  
 DB 117 GKIIIFAAVPPHPPFVENISAQAQITMCPHVIIDVRQLEPVLPLPDINRFFHYNQENTPRM 176

QY 181 RLVCMLYTLRTGGTGDSTFVAGRVMTCPSPDNFLFLVPTVQKTRPFTLPLNPLSS 240  
 DB 177 RLVMALYTLRANSGB-DVFTVSCRVLTRPAPDFEFTFLVPTVESKTKPFTLPLTLGE 235

QY 241 LNSRAPLPISGMIGISPDNVQSFQNGRCTLDGLRVGTTPVLSHVAKIRGT----- 293  
 DB 236 LNSRFPAAIDMLYTDPNESIVVQPNQGRCTLDGLTQGTQVPTQICAFRGLILISQAR 295

QY 294 -----SNGTVINLTDLGTPFHPFEG-PAPIGFDPDLGGCDWHINMTQFCHSQ 340  
 DB 296 AADSTDSQARHNLHVQVKNLDGTQYDFTDDIPAVLGAIDFKG-----TVFGVASQ 348

QY 341 T-----QYDVTDPDTFVPHLGSIQANGISGNY-----IGV 372  
 DB 349 RDVSGQEQGHYATRAHEAHIDITDPKYAPKLTILIKS-GSDDFNTNQPIRPTVGMGD 407

QY 373 LSWVSPSHSGSQVDLWKIPNYGSSITEATHLAPSVYPPGCEVLVFFWMSKIPGAYS 432  
 DB 408 NNRQ-----WELPDYSGRLTNLNNLAPAVSPFPGERILFFRSIVPSAGYG 455

QY 433 ---LPCLLPQYISHLASEQAPTVGEAALLHYDDPTGRTGLGFKAYPDGFLTCVPNGAS 489  
 DB 456 SGYIDCLIPQEWQHFQEAAPSQSAVALVRYNPDTRGNIFEAKLHREGFLTVANCQ-- 513

QY 490 SGPOOLPINGVVFVWSVSRPYQLKPVGTASSAR 523  
 DB 514 NNPVWPNGYFRFEAWGNQFTYLLAPMGSGQGR 547

RESULT 15  
 AAB49710  
 ID AAB49710 standard; protein; 541 AA.  
 XX  
 AC AAB49710;  
 XX  
 DT 04-APR-2001 (first entry)  
 XX  
 DE Small round structured virus protein SEQ ID 11.  
 XX  
 KW Small round structured virus; SRSV; food poisoning.  
 XX  
 OS Small round structured virus.  
 XX  
 PN WO200079280-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 22-JUN-2000; 2000WO-JP004095.  
 XX  
 PR 22-JUN-1999; 99JP-00175928.  
 XX  
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

(DENK-) DENKA SEIKEN KK.  
 Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 WPI; 2001-080848/09.  
 N-PSDB; AAF29151.  
 Kit for the detection and typing of small round-structured virus (SRSV)  
 strains for investigation of food poisoning outbreaks, contains  
 antibodies.  
 Claim 1; Page 64-66; 84pp; Japanese.

This invention relates to a kit for the detection and typing of small  
 round structured virus (SRSV) strains. The kit contains antibodies  
 directed against peptides represented in sequences AAB49700 - AAB49710,  
 which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 used for detecting and typing strains of SRSV in order to prevent the  
 spread of infection and to examine the epidemiology of outbreaks

Sequence 541 AA;

Query Match 35.7%; Score 1005.5; DB 4; Length 541;  
 Best Local Similarity 41.0%; Pred. No. 2.4e-82;  
 Matches 236; Conservative 70; Mismatches 179; Indels 91; Gaps 15;

QY 1 MMASDKATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIINNF 60  
 DB 1 MKMASNDAAAPSSDGAAG--LVPEIN-NEVMLEPVAGASLATPVVGQNIIDPWIRNPF 56

QY 61 VQAPQGEFTISPNNTPGGVLFDSLGPHNLPFLHLSQMYNGWGNMVRIMLAGNAFTA 120  
 DB 57 VQAPAGEFTVSPNPGSEIILDLGPDLPYLAHLMYNGHAGGMEVQIVLAGNAFTA 116

QY 121 GKIIIVCIPGFGSHNLTIAQATLPHVIADVRTLDPIEVLDPEDVRNVLFNHNDROOQM 180  
 DB 117 GKIIIFAAVPPHPPFVENISAQAQITMCPHVIIDVRQLEPVLPLPDINRFFHYNQNDPKL 176

QY 181 RLVCMLYTLRTGGTGDSTFVAGRVMTCPSPDNFLFLVPTVQKTRPFTLPLNPLSS 240  
 DB 177 RLVMALYTLRANSGB-DVFTVSCRVLTRPAPDFEFTFLVPTVESKTKPFTLPLTLGE 236

QY 241 LNSRAPLPISGMIGISPDNVQSFQNGRCTLDGLRVGTTPVLSHVAKIRGT-----TSNGT 297  
 DB 237 MTNSRFPVVDVMTARNENQVQPNQGRVTLGDELLGTTPELLAVNICKPKRGEVIKNGD 296

QY 298 V-----INLTDLGTPFHPFEG-PAPIGFDPDLGGCDWHINMTQFCHSQYD----- 344  
 DB 297 VRSYRMDMEITNDGTDPIDTDPGTGSPDFQ-----ILFGVASQRNKNQNPAT 349

QY 345 -----VDTTPTFVPHLGSIQANGISGNYI---GVLSWVSP----- 378  
 DB 350 RAHEAIINTGGDHLCPQISS-----SEIYLTSPNLRCTNPQPLPOSGLRGTILRSD 402

QY 379 -----PSHPSGSQVDLWKIPNYGSSITEATHLAPSVYPPGCEVLVFFWMSKIPG 427  
 DB 403 NGCHDMVGTSTFTTWPQ--QMRCSRGSNCCSGHRYP--VPVMMNRVTMVLVSHKSG 458

QY 428 PGAYSLPCLLPQYISHLASEQAPTVGEAALLHYDDPTGRTGLGFKAYPDGFLTCVPNG 487  
 DB 459 FSTSTRK--LPQNL-----RWPLIRFINPDTGRVLFEARLHKQGFITVAHTG 504

QY 488 ASSGPOOLPINGVVFVWSVSRPYQLKPVGTASSAR 523  
 DB 505 --DNPIVMPNGYFRFEAWGNQFTYLLAPMGSGQGR 538

Search completed: June 1, 2004, 13:45:46  
 Job time : 47.5302 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.6275 Seconds  
(without alignments)  
2166.837 Million cell updates/sec

Title: US-09-926-799-2  
Perfect score: 2819  
Sequence: 1 MMASKDATSSVDGASGAGQ.....YOLKPVGTASSARGRLGLRR 530

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA\*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2793	99.1	530	US-08-486-049-3	Sequence 3, Appli
2	244.5	8.7	668	US-09-617-594A-4	Sequence 4, Appli
3	239	8.5	623	US-09-590-020-2	Sequence 2, Appli
4	239	8.5	626	US-09-590-020-7	Sequence 7, Appli
5	238	8.4	669	US-09-617-594A-2	Sequence 2, Appli
6	236	8.4	623	US-09-590-020-4	Sequence 4, Appli
7	225	8.0	622	US-09-590-020-6	Sequence 6, Appli
8	129	4.6	2206	US-07-852-260-2	Sequence 2, Appli
9	129	4.6	2206	US-08-461-503-2	Sequence 2, Appli
10	129	4.6	2206	US-08-465-250-2	Sequence 2, Appli
11	127	4.5	517	US-08-689-421-33	Sequence 33, Appli
12	127	4.5	517	US-09-389-528-33	Sequence 33, Appli
13	127	4.5	517	US-09-181-827A-33	Sequence 33, Appli
14	124.5	4.4	3892	US-09-328-352-5503	Sequence 5503, Ap
15	123.5	4.4	2183	US-08-746-111-5	Sequence 5, Appli
16	107.5	3.8	1048	US-09-171-699-10	Sequence 10, Appli
17	106	3.8	20	US-08-973-961-5	Sequence 5, Appli
18	106	3.8	20	US-09-705-621-5	Sequence 5, Appli
19	106	3.8	1042	US-08-928-361B-11	Sequence 11, Appli
20	106	3.8	1042	US-09-588-995A-11	Sequence 11, Appli
21	106	3.8	1837	US-08-928-361B-5	Sequence 5, Appli
22	106	3.8	1837	US-09-588-995A-5	Sequence 5, Appli
23	105	3.7	1043	US-08-928-361B-30	Sequence 30, Appli
24	105	3.7	1721	US-08-700-651-5	Sequence 5, Appli
25	105	3.7	1721	US-08-928-361B-6	Sequence 6, Appli
26	105	3.7	1721	US-09-588-995A-6	Sequence 6, Appli
27	105	3.7	2736	US-09-252-991A-30227	Sequence 30227, A

28 102.5 3.6 770 4 US-09-489-039A-7872 Sequence 7872, Ap  
29 102 3.6 539 2 US-08-808-931-16 Sequence 16, Appl  
30 102 3.6 539 3 US-08-808-323-16 Sequence 16, Appl  
31 102 3.6 539 3 US-09-050-603A-16 Sequence 16, Appl  
32 102 3.6 539 3 US-09-102-420B-16 Sequence 16, Appl  
33 102 3.6 539 4 US-09-497-698-16 Sequence 16, Appl  
34 101.5 3.6 1234 2 US-08-317-310A-15 Sequence 15, Appl  
35 101.5 3.6 1234 5 PCT-US95-13041-15 Sequence 15, Appl  
36 100.5 3.6 2318 3 US-09-091-219-24 Sequence 24, Appl  
37 100.5 3.6 2318 4 US-09-660-541-24 Sequence 24, Appl  
38 99 3.5 512 1 US-08-462-484-6 Sequence 6, Appl  
39 99 3.5 512 1 US-08-441-147-6 Sequence 6, Appl  
40 99 3.5 512 5 PCT-US95-07536-5 Sequence 6, Appl  
41 98.5 3.5 803 4 US-09-543-681A-5399 Sequence 5399, Ap  
42 97 3.4 469 4 US-09-479-645A-6 Sequence 6, Appl  
43 97 3.4 469 4 US-09-479-645A-8 Sequence 8, Appl  
44 97 3.4 2442 4 US-09-514-247A-10 Sequence 10, Appl  
45 96 3.4 2556 1 US-08-185-432-17 Sequence 17, Appl

ALIGNMENTS

RESULT 1  
US-08-486-049-3  
; Sequence 3, Application US/08486049  
; Patent No. 6572862  
; GENERAL INFORMATION:  
; APPLICANT: Bates, Mary K  
; APPLICANT: Jiang, Xi  
; APPLICANT: Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE:  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-486-049-3

Query Match 99.1%; Score 2793; DB 4; Length 530;  
Best Local Similarity 98.7%; Pred. No. 1.9e-253;  
Matches 523; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVNPIDPWIINNF 60  
DB 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVNPIDPWIINNF 60

QY 61 VOAPQGETTISNNTPGGVLDLSLGLPHLNPFLHLSQMYNGWGNMRVIMLAGNAFTA 120  
DB 61 VOAPQGETTISNNTPGGVLDLSLGLPHLNPFLHLSQMYNGWGNMRVIMLAGNAFTA 120  
QY 121 GKIIIVSCIPPPGGSHNLTTIAQATLPPHVIADVRTLDPIEVPLEDRVNLFNHNDNRQOQTM 180  
DB 121 GKIIIVSCIPPPGGSHNLTTIAQATLPPHVIADVRTLDPIEVPLEDRVNLFNHNDNRQOQTM 180  
QY 181 RLVCMLYTPRLRTGGTGDGDFVAVAGVMTCPSPDFNPLFLVPTVQKTRPFTLNLPLSS 240  
DB 181 RLVCMLYTPRLRTGGTGDGDFVAVAGVMTCPSPDFNPLFLVPTVQKTRPFTLNLPLSS 240  
QY 241 LNSRAPLPISMGISPDNVQSVQFONGRCTLDGLVGTTPVSLSHVAKIRGTSNGTVIN 300  
DB 241 LNSRAPLPISMGISPDNVQSVQFONGRCTLDGLVGTTPVSLSHVAKIRGTSNGTVIN 300  
QY 301 LTELDTGTPHPFPGPAPIGFDDLGCDMHNMTQFGHSSQOQYDVTTPDTFVPHLGSIQ 360  
DB 301 LTELDTGTPHPFPGPAPIGFDDLGCDMHNMTQFGHSSQOQYDVTTPDTFVPHLGSIQ 360  
QY 361 ANGIGSNVIGVLSWSPSPHSGSOVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420  
DB 361 ANGIGSNVIGVLSWSPSPHSGSOVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVF 420  
QY 421 FMSKIPGAYSLPCLLPQEIYISHLASEQAPTVGEAALLHYVDPDTRTLGFBKAYPDGF 480  
DB 421 FMSKIPGAYSLPCLLPQEIYISHLASEQAPTVGEAALLHYVDPDTRTLGFBKAYPDGF 480  
QY 481 LTCVPNGASSGPQOLPINGVVFVSVWSRYQLKPVGTASSARGRLGLAR 530  
DB 481 LTCVPNGASSGPQOLPINGVVFVSVWSRYQLKPVGTASSARGRLGLAR 530

RESULT 2  
US-09-617-594A-4  
; Sequence 4, Application US/09617594A  
; Patent No. 6541458  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALCIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.1  
; CURRENT APPLICATION NUMBER: US/09/617,594A  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 668  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-617-594A-4

Query Match 8.7%; Score 244.5; DB 4; Length 668;  
Best Local Similarity 32.2%; Pred. No. 4.3e-14;  
Matches 69; Conservative 27; Mismatches 87; Indels 31; Gaps 8;  
QY 19 GOLVENVASDPLAMPVAGSTAVATAGVNPIDPWIINNFPVQAPQGEFTI-----SP 72  
DB 136 GTLVGVIAEPNAQMSAVAD----VATGKSDV--SEW-----EAFHFHTSVNWST 180  
QY 73 NNTPGGVLDLSLGLPHLNPFLHLSQMYNGWGNMRVIMLAGNAFTAAGKIIVSCIPP 132  
DB 181 SETQKILFKQSLGFLNPLPYLTHLAKLYVWSSIEVRFSGSGVFGKLAIVVPPGI 240  
QY 133 GSHNLTTIAQATLPPHVIADVRTLDPIEVPLEDRVNLFNHNDNRQOQTMRLVCMVLY 188  
DB 241 DPVQST--SMLQVPHVLFDAQVPEVIFIPDLRNSLYHLS- DTDTSLSVIMINDLIN 297

QY 189 PLRTGGTGDGDFVAVAGVMTCPSPDFNPLFLVPP 222  
DB 298 PYANDSNSGCIVT---VETKPGDPFKFHLKAPP 328

RESULT 3  
US-09-590-020-2  
; Sequence 2, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Aivars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-2

Query Match 8.5%; Score 239; DB 4; Length 623;  
Best Local Similarity 26.3%; Pred. No. 1.3e-13;  
Matches 74; Conservative 43; Mismatches 98; Indels 66; Gaps 9;  
QY 71 SPNTPGGVLDLSLGLPHLNPFLHLSQMYNGWGNMRVIMLAGNAFTAAGKIIVSCIPP 130  
DB 134 STSETQKILFKQSLGFLNPLPYLTHLAKLYVWSSIEVRFSGSGVFGKLAIVVPP 193  
QY 131 GFGSHNLTTIAQAT---LPPHVIADVRTLDPIEVPLEDRVNLFNHNDNRQOQTMRLVCMVLY 187  
DB 194 G-----IEPVQSTMLQYPHVLFDAQVPEVIFIPDLRNSLYHLS- DTDTSLSVIMVY 247  
QY 188 TPLRT--GGTGDGDFVAVAGVMTCPSPDFNPLFLVPP----- 222  
DB 248 NDLPINPYANDTNSGCIV-TVETKPGDPFKFHLKAPPQSLTHGSPVSDLPKSSSLWIG 306  
QY 223 -----TVEQKTRPFTLNLPLSLSLNSRAPLPISGNGISPDN 259  
DB 307 NRHMSDITDFTIRPFVQANRHFDFNQETAGWSTPRFRPITITVSESNMKLGIGVATDY 366  
QY 260 -VQSVQFONGRCTLDGLVGTTPVSLSHVAKIRGTSNGTVI 299  
DB 367 IVEGIPDGPWPTTPEQL---TPAGIYSIT-----ASNGTVI 400

RESULT 4  
US-09-590-020-7  
; Sequence 7, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Aivars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 626



```

; TYPE: PR1
; ORGANISM: Feline calicivirus
US-09-590-020-7

Query Match
Best Local Similarity 8.5%; Score 239; DB 4; Length 626;
Matches 71; Conservative 30; Mismatches 90; Indels 30; Gaps 8;

QY 13 DGASGA---GQIVPEVNASDPLAMPVAGSSAVATAGQVNPIDPWIIINNFVQAPQGEPT 69
Db 82 DGSITAEQGTWGGVIA-----BPSAQMSAADMATCKSVDSW-----EAPFS 126

QY 70 I-----SPNTPGGVLDLSLGPPLNPLHLSQMYNGWGNMVRIMLAGNATAGKI 123
Db 127 FHTSVNWSSTQKILFKQSLGPLNPLHLSQMYNGWGNMVRIMLAGNATAGKI 186

QY 124 IVSCIPPGGSHNLTAQATLFPVHVIADVRLDPIEVPLEDVNRVLFHNDRNQOTMRLV 183
Db 187 AAIVVPPGVDPVQST--SMLQYPHVLFDAQVPEVIFCLPDLRSTLYHLMS-DTDTTSLV 243

QY 184 CMLYTPLR--GGGTGDSFVAVAGRVMTCPSPDNFLFLVPP 222
Db 244 INVYNDLINPYANDTNSSGCIV-TVETKGPDPFKFHLKPP 283

RESULT 5
US-09-617-594A-2
; Sequence 2, Application US/09617594A
; Patent No. 6541458
; GENERAL INFORMATION:
; APPLICANT: Audomnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PR1
; ORGANISM: Feline calicivirus
US-09-617-594A-2

Query Match
Best Local Similarity 8.4%; Score 238; DB 4; Length 669;
Matches 67; Conservative 32; Mismatches 92; Indels 38; Gaps 8;

QY 7 DATSSV---DGASGAGQLVPEVNASDPLAMPVAGSSAVATAGQVNPIDPWIIINNFVQA 63
Db 126 DGDSSITTEQGLTGGVIAEBSAQMTAADAATGKSD-----SEW----- 167

QY 64 PQGEFTI-----SPNTPGGVLDLSLGPPLNPLHLSQMYNGWGNMVRIMLAGNA 117
Db 168 -ESFFSHTSVNWSSTQKILFKQSLGPLNPLHLSQMYNGWGNMVRIMLAGNA 226

QY 118 FTAGKIIVSCIPPGGSHNLTAQATLFPVHVIADVRLDPIEVPLEDVNRVLFHNDRNQ 177
Db 227 VFGGKLAIVVPPGVDPVQST--SMLQYPHVLFDAQVPEVIFCLPDLRSTLYHLMS-DT 283

QY 178 QTMRLVCMY-----TPLRTGGGTGDSFVAVAGRVMTCPSPDNFLFLVPP 222
Db 284 DTTSLVIMVYNDLINPYANDTNSSGCIVT---VETKGPDPFKFHLKPP 329

RESULT 6
US-09-590-020-4
; Sequence 4, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PR1
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match
Best Local Similarity 8.4%; Score 236; DB 4; Length 623;
Matches 56; Conservative 27; Mismatches 62; Indels 12; Gaps 5;

QY 71 SPNTPGGVLDLSLGPPLNPLHLSQMYNGWGNMVRIMLAGNATAGKIIVSCIPP 130
Db 134 STSETQKILFKQSLGPLNPLHLSQMYNGWGNMVRIMLAGNATAGKIIVSCIPP 193

QY 131 GFGSHNLTAQAT---LFPVHVIADVRLDPIEVPLEDVNRVLFHNDRNQOTMRLVCMY 187
Db 194 G-----IEPVQSTSMLOQPHVLFDAQVPEVIFCLPDLRSTLYHLMS-DTDTTSLVIMVY 247

QY 188 TPLRT--GGGTGDSFVAVAGRVMTCPSPDNFLFLVPP 222
Db 248 NDLINPYANDTNSSGCIV-TVETKGPDPFKFHLKPP 283

RESULT 7
US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PR1
; ORGANISM: Feline calicivirus
US-09-590-020-6

Query Match
Best Local Similarity 8.0%; Score 235; DB 4; Length 622;
Matches 55; Conservative 27; Mismatches 63; Indels 12; Gaps 5;

QY 71 SPNTPGGVLDLSLGPPLNPLHLSQMYNGWGNMVRIMLAGNATAGKIIVSCIPP 130
Db 134 STSETQKILFKQSLGPLNPLHLSQMYNGWGNMVRIMLAGNATAGKIIVSCIPP 193

QY 131 GFGSHNLTAQAT---LFPVHVIADVRLDPIEVPLEDVNRVLFHNDRNQOTMRLVCMY 187
Db 194 G-----IEPVQSTSMLOQPHVLFDAQVPEVIFCLPDLRSTLYHLMS-DTDTTSLVIMVY 247

QY 188 TPLRT--GGGTGDSFVAVAGRVMTCPSPDNFLFLVPP 222

```

Db 248 NDLINPYANDTNSGCIIV-TVETKGPDPFKFHLKPP 283

Db 630 PLAPSDTVQTRHVQVRRSRSESTIESFFARGACVAIIIVDN-----EQPTTRAKLFAM 683

Qy 318 --IGFPDLGGCDWHINMTQFGHSSOTQYDQVDTPTDFVPHLGSIQANGIGSGNYIGVLSW 375

Db 684 WRITYKDT--VQLRRKLEFFYS---RFDMEF---TFVVTANFTNANGHALNQVQIMY 735

Qy 376 VSPSPHSGQVLDLWKIPNYGSSITEATHLAPSVY 410

Db 736 I-PPGAPT-----PKSWDDYTQWTSNPSIF 760

RESULT 9

US-08-461-503-2

; Sequence 2, Application US/08461503

; Patent No. 5834302

; GENERAL INFORMATION:

; APPLICANT: Racaniello, Vincent

; APPLICANT: Tatem, Joanne M.

; APPLICANT: Weeks-Levy, Carolyn L.

; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES

; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES

; TITLE OF INVENTION: FROM CDNA

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham

; STREET: 1185 Avenue of the Americas

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10112

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/461,503

; FILING DATE: 5-JUN-1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 278-0400

; TELEFAX: (212) 391-0525

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2206 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-461-503-2

Query Match 4.6%; Score 129; DB 2; Length 2206;

Best Local Similarity 21.1%; Pred. No. 0.02;

Matches 96; Conservative 59; Mismatches 196; Indels 104; Gaps 20;

Qy 17 GAGQLVPEVNASDPLAMPVAGSSTAVATAGQVN-----PIDPWIINNFVQAPQG---- 66

Db 349 GSNQYLTSNDHQSPCAI-PEFDVTPPIDIPGEVKNMELAEIDTMIPLNLESTKRNMTDM 407

Qy 67 -EFTISPNNTGGVLFDSLGLPHLNPFLLH-----LSQWYNGWGNMVRIMLAGNAFTA 120

Db 408 YRVTLSDSADLSQPLICLSLSPAFDPRLSHTMLGEVLNYTHWAGSLKFTFLFCGSMWAT 467

Qy 121 GKIIIVSCIPPGFGSHNLTIQAATLFPFHVIADVRTLDPIEVLPLEVDRVNLFHNDRNQOT- 179

Db 468 GKILVAYAPPG-AQPPTSKEAMLGTHVWDLGLQSSCTMVVPVWISNVTYRQTQDSFTE 526

Qy 180 -----MRLVCMLYTPTLRTGGTGDSSFVAGRVMTCPSPDENFLFL-----V 220

Db 527 GGYISMFYQTRIVVPLSTP-----KSMMLGFSVAC--NDFSRLLRDTHHSQSAL 576

Db 248 NDLINPYANDTNSGCIIV-TVETKGPDPFKFHLKPP 283

Db 630 PLAPSDTVQTRHVQVRRSRSESTIESFFARGACVAIIIVDN-----EQPTTRAKLFAM 683

Qy 318 --IGFPDLGGCDWHINMTQFGHSSOTQYDQVDTPTDFVPHLGSIQANGIGSGNYIGVLSW 375

Db 684 WRITYKDT--VQLRRKLEFFYS---RFDMEF---TFVVTANFTNANGHALNQVQIMY 735

Qy 376 VSPSPHSGQVLDLWKIPNYGSSITEATHLAPSVY 410

Db 736 I-PPGAPT-----PKSWDDYTQWTSNPSIF 760

RESULT 8

US-07-852-260-2

; Sequence 2, Application US/07852260

; Patent No. 5525715

; GENERAL INFORMATION:

; APPLICANT: Racaniello, Vincent

; APPLICANT: Tatem, Joanne M.

; APPLICANT: Weeks-Levy, Carolyn L.

; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM

; TITLE OF INVENTION: CDNA

; NUMBER OF SEQUENCES: 9

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Cooper & Dunham

; STREET: 30 Rockefeller Plaza

; CITY: New York

; STATE: New York

; COUNTRY: U.S.A.

; ZIP: 10112

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/07/852,260

; FILING DATE: 19920619

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: White, John P.

; REGISTRATION NUMBER: 28,678

; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 977-9550

; TELEFAX: (212) 664-0525

; INFORMATION FOR SEQ ID NO: 2:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2206 amino acids

; TYPE: AMINO ACID

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-07-852-260-2

Query Match 4.6%; Score 129; DB 1; Length 2206;

Best Local Similarity 21.1%; Pred. No. 0.02;

Matches 96; Conservative 59; Mismatches 196; Indels 104; Gaps 20;

Qy 17 GAGQLVPEVNASDPLAMPVAGSSTAVATAGQVN-----PIDPWIINNFVQAPQG---- 66

Db 349 GSNQYLTSNDHQSPCAI-PEFDVTPPIDIPGEVKNMELAEIDTMIPLNLESTKRNMTDM 407

Qy 67 -EFTISPNNTGGVLFDSLGLPHLNPFLLH-----LSQWYNGWGNMVRIMLAGNAFTA 120

Db 408 YRVTLSDSADLSQPLICLSLSPAFDPRLSHTMLGEVLNYTHWAGSLKFTFLFCGSMWAT 467

Qy 121 GKIIIVSCIPPGFGSHNLTIQAATLFPFHVIADVRTLDPIEVLPLEVDRVNLFHNDRNQOT- 179

Db 468 GKILVAYAPPG-AQPPTSKEAMLGTHVWDLGLQSSCTMVVPVWISNVTYRQTQDSFTE 526

Qy 180 -----MRLVCMLYTPTLRTGGTGDSSFVAGRVMTCPSPDENFLFL-----V 220

Db 527 GGYISMFYQTRIVVPLSTP-----KSMMLGFSVAC--NDFSRLLRDTHHSQSAL 576

Qy 221 PPTVQKTRPFTLNLPLS-----SLNSRAPLPISGMISPDNVQSVQFQNGRCLTDG 274

Db 577 PGIGEDLTSVAQAGALTSLPKQDLSLPDKA-----SGPAHSKEVPALTAVETG---ATN 629

Qy 275 RLVGTTPTVSLSHVAKIRGTSNGTV-----INLTEDGTGTFPHFEGRAP----- 317

QY 221 PPTVEQKTRPFTLPNPLS-----SLNSRAPLPISGMISPDNVQSVQFONGRCTLG 274  
DB 577 PGQIEDLTSEVAQAGALTSLPKQDLSLPTKA-----SGPAHSKEVPALTA VETG---ATN 629  
QY 275 RLVGTPPVSLSHVAKIRGTSNGTV-----INLTDLGTPFPFPFGPAP----- 317  
DB 630 PLAPSDTVQTRHVQRRSRSESTISFFARGACVAIIEDVN-----EQPTTTRAKLFAM 683  
QY 318 --IGPPDLGGCDWHINMTQFGHSSQTDVDTPTTFVPHLGSIOANGIGSGNYIGVLSW 375  
DB 684 WRITYKDT--VQLRRKLEFFTYS---RFDMEF---TFVVTANFTNANNCHALNQVQIMY 735  
QY 376 VSPPSHPSGSQVDLWKIPNYGSSITEATHLAPSVY 410  
DB 736 I-PPGAPT-----PKSWDDYTQWTSNPSIF 760

## RESULT 10

US-08-465-250-2  
; Sequence 2, Application US/08465250  
; Patent No. 6136570  
; GENERAL INFORMATION:  
; APPLICANT: Racaniello, Vincent  
; APPLICANT: Tatem, Joanne M.  
; APPLICANT: Weeks-Levy, Carolyn L.  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM  
; TITLE OF INVENTION: CDNA  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release 1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,250  
; FILING DATE: 6-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2206 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-465-250-2

Query Match 4.6%; Score 129; DB 3; Length 2206;  
Best Local Similarity 21.1%; Pred. No. 0.02;  
Matches 96; Conservative 59; Mismatches 196; Indels 104; Gaps 20;

QY 17 GAGQLVPEVNASDPLAMPDVAGSSTAVATAGVN-----PIDPWIINNFVQAPQG----- 66  
DB 349 GSNQVLTSDNHQSPCAI-PEFDVTPPIDPGEVQKMWELAEIDTWIPLNLESTKENTMDM 407  
QY 67 -EFTISPNNTPGGVFLDLSLGLHNPFLH-----LSQWYNGWGNMVRIMLAGNAFTA 120  
DB 408 YRVLTSDSADLSQPLCLSLSPAFDRLSHTMLGEVLNYYTHWAGSLKFTFLFCGSMWAT 467

QY 121 GKLIYSCIPPGSGSHNLTIQAATLPPHVIADVRTLDPIEVLPLEDVRNVLPHNDRNQQT- 179  
DB 468 GKILVAYAPPG-AQPTSRKEAMLGTHVWDLGLOSSCTMVPVWISNTYRTQTDQSFT 526  
QY 180 -----MRLVCMLYTPLRTGGGTGDSFVVAGRVMTCPSPDFNLFEL-----V 220  
DB 527 GGYISMPFQTRIVVPLSTP-----KSMMLGFVSAC--NDFSVRLLRDTTHISQSAL 576  
QY 221 PPTVEQKTRPFTLPNPLS-----SLNSRAPLPISGMISPDNVQSVQFONGRCTLG 274  
DB 577 PGQIEDLTSEVAQAGALTSLPKQDLSLPTKA-----SGPAHSKEVPALTA VETG---ATN 629  
QY 275 RLVGTPPVSLSHVAKIRGTSNGTV-----INLTDLGTPFPFPFGPAP----- 317  
DB 630 PLAPSDTVQTRHVQRRSRSESTISFFARGACVAIIEDVN-----EQPTTTRAKLFAM 683  
QY 318 --IGPPDLGGCDWHINMTQFGHSSQTDVDTPTTFVPHLGSIOANGIGSGNYIGVLSW 375  
DB 684 WRITYKDT--VQLRRKLEFFTYS---RFDMEF---TFVVTANFTNANNCHALNQVQIMY 735  
QY 376 VSPPSHPSGSQVDLWKIPNYGSSITEATHLAPSVY 410  
DB 736 I-PPGAPT-----PKSWDDYTQWTSNPSIF 760

## RESULT 11

US-08-689-421-33  
; Sequence 33, Application US/08689421  
; Patent No. 6008029  
; GENERAL INFORMATION:  
; APPLICANT: Yaver, Debbie S.  
; APPLICANT: Brown, Kimberley M.  
; APPLICANT: Kauppinen, Sakari  
; APPLICANT: Halkier, Torben P.  
; TITLE OF INVENTION: PURIFIED COPRINUS LACCASES AND NUCLEIC ACIDS  
; TITLE OF INVENTION: ENCODING SAME  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NO. 60080290 No. 6008029disk of No. 6008029th America, Inc.  
; STREET: 405 Lexington Avenue, 64th Floor  
; CITY: New York  
; STATE: New York  
; COUNTRY: United States of America  
; ZIP: 10174-6401

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/689,421  
; FILING DATE: 9-AUG-1996  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Harrington, James J.  
; REGISTRATION NUMBER: 38,711  
; REFERENCE/DOCKET NUMBER: 4554.204-WO  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-867-0123  
; TELEFAX: 212-878-9655  
; INFORMATION FOR SEQ ID NO: 33:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 517 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-689-421-33

Query Match 4.5%; Score 127; DB 3; Length 517;  
Best Local Similarity 23.8%; Pred. No. 0.003;  
Matches 124; Conservative 59; Mismatches 200; Indels 138; Gaps 30;

QY 40 STAVATA-----GOVNPIDPWIINNFVQAPQGBFTIS-----PN-NTPGGVL----- 80

TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-389-528-33

Query Match 4.5%; Score 127; DB 3; Length 517;  
Best Local Similarity 23.8%; Pred. No. 0.003;  
Matches 124; Conservative 59; Mismatches 200; Indels 139; Gaps 30;

5 ATALATSLPFLVGLGAGSTNLWANKVAPDG-FRSRAVLGATQTPVQPGFVIQGNK 63  
81 ---FDLSLGPLH-NPFLHLSONYGNWVRVIMLAGNAFTAGKIIVSCIPPGFGSHN 136  
64 NSFFAINVIDALTDPTMLRTTSH--WHG-----MFQGTAWADGPGAGVTCQIPSPGHSHF 116  
137 LTIAQA-----TLFPHVIADVRLDPIE-----VPLEDVNRVLFHNDRNQOTMRLVCM 185  
117 LYKFOALNOAGTFTWYSHSHESQYCDLRCAMVVYDVPDHRNL--YDIDNEATITLADW 174  
186 LYTLPLRTGG--GTGDSFVWAGRYMTCPSPDFNLFVLPPTVEQKTR-----PFTL----- 233  
175 YHVPAAPSAGLVPTPDSTLNGKRYAGGP-----TVPLAVISVTRNRRYRFLVLSLSCD 228  
234 PNLPLSSLSNRAPLPISGMGI SPDNVQSVQF-----QNGRGT 271  
229 PNYVFSIDGHTMTVIEVDGVNVQPLVVDLSIQIFAGORYSFVLNANRPVGNVWRANPNIG 288  
272 LDGRLVGTTPVSLSHVAKIRGTSNGTVINLTLDGTFP-----HPFEGPAPIGPPD 322  
289 TTGFVGVNSALRYV---GASN-TDPTTQ---TPFSNPLETLNHLPLTNPAAPGLPT 340  
323 LGGCDWHINM-TQGHSSQTYQYDVTPTDF-----VPHLGSIQANG-----IGSGN-YI 370  
341 PGGVDVAIINLTVPDFSSLT-FSVNGA--TFHQPPVPVLLQIMSGAQTAAQLLPSSGVYV 397  
371 GVLSSWSPSPSHSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFG-----QNGRGT 415  
398 LPRNKVIELSMGPGSTGSPHFLHGHEFAVVRSGSSTY--NFANPVRDVRVVSAGVAGD 455  
416 EVLVFFMSKIPGPGAYSPLCLLPQEIYSHLA---SEQAPTV 453  
456 NVTIIRFTDNEGP--WILHCHIDWHLVGLAVVFAEDAPT 494

RESULT 12  
US-09-389-528-33  
Sequence 33, Application US/09389528  
Patent No. 6207430  
GENERAL INFORMATION:  
APPLICANT: Yaver, Debbie S.  
APPLICANT: Brown, Kimberley M.  
APPLICANT: Kauppinen, Sakari  
APPLICANT: Halkier, Torben P.  
TITLE OF INVENTION: PURIFIED COPRINUS LACCASES AND NUCLEIC ACIDS  
TITLE OF INVENTION: ENCODING SAME  
NUMBER OF SEQUENCES: 36  
CORRESPONDENCE ADDRESS:  
ADDRESS: No. 6207430 No. 6207430disk of No. 6207430th America, Inc.  
STREET: 405 Lexington Avenue, 64th Floor  
CITY: New York  
STATE: New York  
COUNTRY: United States of America  
ZIP: 10174-6401  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/389,528  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/689,421  
FILING DATE: 9-AUG-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Harrington, James J.  
REGISTRATION NUMBER: 38,711  
REFERENCE/DOCKET NUMBER: 4554.204-WO  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-867-0123

TELEFAX: 212-878-9655  
INFORMATION FOR SEQ ID NO: 33:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 517 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-389-528-33

Query Match 4.5%; Score 127; DB 3; Length 517;  
Best Local Similarity 23.8%; Pred. No. 0.003;  
Matches 124; Conservative 59; Mismatches 200; Indels 139; Gaps 30;

40 STAVATA-----GOVNPIDPMIINNFWQAPQGEFTIS-----PN-NTPGGVL----- 80  
5 ATALATSLPFLVGLGAGSTNLWANKVAPDG-FRSRAVLGATQTPVQPGFVIQGNK 63  
81 ---FDLSLGPLH-NPFLHLSONYGNWVRVIMLAGNAFTAGKIIVSCIPPGFGSHN 136  
64 NSFFAINVIDALTDPTMLRTTSH--WHG-----MFQGTAWADGPGAGVTCQIPSPGHSHF 116  
137 LTIAQA-----TLFPHVIADVRLDPIE-----VPLEDVNRVLFHNDRNQOTMRLVCM 185  
117 LYKFOALNOAGTFTWYSHSHESQYCDLRCAMVVYDVPDHRNL--YDIDNEATITLADW 174  
186 LYTLPLRTGG--GTGDSFVWAGRYMTCPSPDFNLFVLPPTVEQKTR-----PFTL----- 233  
175 YHVPAAPSAGLVPTPDSTLNGKRYAGGP-----TVPLAVISVTRNRRYRFLVLSLSCD 228  
234 PNLPLSSLSNRAPLPISGMGI SPDNVQSVQF-----QNGRGT 271  
229 PNYVFSIDGHTMTVIEVDGVNVQPLVVDLSIQIFAGORYSFVLNANRPVGNVWRANPNIG 288  
272 LDGRLVGTTPVSLSHVAKIRGTSNGTVINLTLDGTFP-----HPFEGPAPIGPPD 322  
289 TTGFVGVNSALRYV---GASN-TDPTTQ---TPFSNPLETLNHLPLTNPAAPGLPT 340  
323 LGGCDWHINM-TQGHSSQTYQYDVTPTDF-----VPHLGSIQANG-----IGSGN-YI 370  
341 PGGVDVAIINLTVPDFSSLT-FSVNGA--TFHQPPVPVLLQIMSGAQTAAQLLPSSGVYV 397  
371 GVLSSWSPSPSHSGSQVDLWKIPNYGSSITEATHLAPSVYPPGFG-----QNGRGT 415  
398 LPRNKVIELSMGPGSTGSPHFLHGHEFAVVRSGSSTY--NFANPVRDVRVVSAGVAGD 455  
416 EVLVFFMSKIPGPGAYSPLCLLPQEIYSHLA---SEQAPTV 453  
456 NVTIIRFTDNEGP--WILHCHIDWHLVGLAVVFAEDAPT 494

RESULT 13  
US-09-181-827A-33  
Sequence 33, Application US/09181827A  
Patent No. 6242232  
GENERAL INFORMATION:  
APPLICANT: Yaver, Debbie S.  
APPLICANT: Brown, Kimberley M.  
APPLICANT: Kauppinen, Sakari  
APPLICANT: Halkier, Torben P.  
TITLE OF INVENTION: Purified Coprinus Laccases And Nucleic  
TITLE OF INVENTION: Acids Encoding Same  
FILE REFERENCE: 4554.200-US  
CURRENT APPLICATION NUMBER: US/09/181,827A  
CURRENT FILING DATE: 1998-10-28  
PRIOR APPLICATION NUMBER: 60/002,800  
PRIOR FILING DATE: 1995-08-25  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 33  
LENGTH: 517  
TYPE: PRT  
ORGANISM: Coprinus cinereus  
US-09-181-827A-33

Query Match 4.5%; Score 127; DB 3; Length 517;  
Best Local Similarity 23.8%; Pred. No. 0.003;  
Matches 124; Conservative 59; Mismatches 200; Indels 136; Gaps 30;

QY 40 STAVATA-----GQVNPIDPMLNNFVQAPQGEFTIS-----PN-NTGGVL----- 80  
DB 5 ATALATSLLPFLVGLGAPSTNLVANKVIAPDG-FSRSAVLAGATQPTVQFPVPGVIGNK 63  
QY 81 ----FOLSLGPHL-NPFLHLSQMYNGWGNMVRIMLAGNAFTAGKIIIVSCIPPGFGSHN 136  
DB 64 NSFFAINVIDALTDPTMLRTTSH--WHG-----MFQGTAWADGAGVTCQPIPGHSF 116  
QY 137 LIAQA-----TLFPHVIAVRLTDPIE-----VPLEDRVNLVFNHNDNRNQOTWRLVCM 185  
DB 117 LYKFOALNOAGTFWYHSHHESQYCDGLRGAMVYDVPDPRNL--YDIDNEATIITLADW 174  
QY 186 LYTLERTGG--GTGGSFVAGRVMTCPSPDFNLEFLVPTVEQKTR-----PTL----- 233  
DB 175 YXVPAPSAGLVPTDPTSLINGRYAGGP-----TVPLAVISVTRNRRYRFLVLSLSD 228  
QY 234 PNLPGLSSLSNRAPIPIGMSGISPONVQSVQF-----QNGRCT 271  
DB 229 PNYVFSIDGHTVTVIEVDGVNVQPLVDSIQIFAGORYSFVLNANRPVGNVYVRANPNIG 288  
QY 272 LDGRLVGTTPVSLSHVAKIRGTSNGTVINLTLDGTPF-----HPFEGPAPIGPPD 322  
DB 289 TTGFVGGVNSAILRYV---GASN-TDPTTQ---TPFSNPLETLNHLPLTPAAPGLPT 340  
QY 323 LGGCDWHNM-TQGHSSQOTQDVDTPTDF-----VPHLSIQANG-----IGSN-YI 370  
DB 341 PGGVDVAINLNTVDFSSLT-FSVNGA--TFHQPPVPLVLLQIMSGAQTAAQLLPSSGVV 397  
QY 371 GYLSWVSPSPHPSGSGQVLDWKIPNYGSSITEATHLAPSVYPPGFG----- 415  
DB 398 LPRNKVIELSMGGSTGSPHPLHGHPEAVVRSAGSTY--NFANPVRDVSAGVAGD 455  
QY 416 EVLVFMSKIPGAGYSLPCLLPQEVYISHLA---SEQAPTV 453  
DB 456 NVTIRFRDTPGPF--WILCHIDHVLVLGLAVVFAEDAPT 494

RESULT 14  
US-09-328-352-5503  
; Sequence 5503, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 5503  
; LENGTH: 3892  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-5503

Query Match 4.4%; Score 124.5; DB 4; Length 3892;  
Best Local Similarity 21.0%; Pred. No. 0.13;  
Matches 125; Conservative 70; Mismatches 243; Indels 157; Gaps 31;

QY 7 DATSSVDGASGAGLV-PRVNASD---PLAMPDVAGSS-TAVATAGVNP-----ID 53  
DB 622 DTATVAGDGSWSVPNPGNLGDVEEATATDPAGNPSLPGTATVDAGPNTDGVNFTVD 681  
QY 54 PWIINNFOAPOGEFTISPNNTPGVLDLSLGPLNLPFLHLSQMYNGWGNMVRIML 113  
DB 682 SVTADNVINASEA-----SGNVTVTGVLKNVPAD-----AANTVVTVVI 720  
QY 114 AGNAFTAGKIIIVSCIPPGFGSHNLTAQATLPPHVIADV-RTLDPIEVLDRVNLVFN 172

DB 721 NGQTYTA-----TVDSSTAGTWTWVSPGSDL-----TADADKTID-AKVTFTDAAGNSSSV 769  
QY 173 NDRNQOTMRLV---CMLYTPLR-----GGT-----GDSFVAVG--RVMTCP 210  
DB 770 NDTQTYTIDTAPDAPVINPVNGTDPITGTAEFGSTVTVYPDGSTTTVVAGPDGTWTV 829  
QY 211 SPDFNLEFLVPTVEQKTRPFLPNLPLSSLSNSRAP-----LPISGMGISPONVQSVQFO 266  
DB 830 NPLGNDGKVTALIA---TDPAGNPSLPGTATVDAGPNTDGVNFTVDSVTADNVINASEA 886  
QY 267 NGRCTLDGRL-----VGTTPVSL-----SHVAKIRGTSNGTVINLTDELDTGTPFPFEGP 315  
DB 887 SGNVTVTGVLKNVPADAANTVTVVINGQTYTATVDSSTAGTWTWVSPGSDLTADADKTID 946  
QY 316 APIGPDPDGGCDWHNMTQFGHSSQOTQDVDTPTDFVPHLGSIQANGIGSGNYIGVLSW 375  
DB 947 AKVTFTDAAGNSSSVNDTH-----TYVDTVAEN-APVLDPINATDPVSGQ----- 991  
QY 376 VSPPS-----HPSGQVDL-----WKIPNYGSSITEATHLAPSVYPPGGEVLVFPMS 423  
DB 992 AEPGSTVTVYTPDGTATVAVAGPDGWSVPNPGNLVDGDTVTATATDPAG-----NT 1043  
QY 424 KIPGCAVSLPCLLPQEVYISH-LASEQAPTVGEAALLHYVDDPTGRTLGEFKA--YP--- 477  
DB 1044 SLPGTGTVSADITAPVVALDDVLTNDSTP-----ALTGTVNDPTATVVNVVDGVDPVAVN 1098  
QY 478 --DGLTCTCPN---CASSGPOQLPINGVVFVSVWSRFRVQLKPKVPTASSARGLG 527  
DB 1099 NGDGTWTLADNLTPLATDGPHTITV-----TATDAAGNAG 1133

RESULT 15  
US-08-746-111-5  
; Sequence 5, Application US/08746111  
; Patent No. 6066778  
; GENERAL INFORMATION:  
; APPLICANT: Ginsburg, David  
; APPLICANT: Cui, Jisong  
; TITLE OF INVENTION: Compositions And Methods For Screening  
; TITLE OF INVENTION: Compounds For Anticoagulant Activity  
; NUMBER OF SEQUENCES: 54  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Medlen & Carroll, LLP  
; STREET: 220 Montgomery Street, Suite 2200  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States of America  
; ZIP: 94104  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION NUMBER: US/08/746,111  
; FILING DATE: 06-NOV-1996  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ingolia, Diane E.  
; REGISTRATION NUMBER: 40,027  
; REFERENCE/DOCKET NUMBER: UM-02536  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 705-8410  
; TELEFAX: (415) 397-8338  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2183 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-746-111-5

```
Query Match      4.4%; Score 123.5; DB 3; Length 2183;
Best Local Similarity 26.8%; Pred. No. 0.063;
Matches 82; Conservative 31; Mismatches 110; Indels 83; Gaps 18;

QY 212 PDFNLFVLVPTVEQKTRPFTLPNLPSSLSNSRAPLPISG-MGISPDNVQSVQFQNGRC 270
Db 1196 PDLQQL-LLSPEDNQKTSPPDLGQVPLSPDDNQKTSPPD--- 1251
QY 271 TLGRLVGTTFVSLSHVAKIRGTSNGTVINLTDLGTFPHPEGPAPIGFPDGLGCDWHI 330
Db 1252 -----LGQVPLSLDDNQKTSPPDLGQV-----PLSPDDNQK-ITSPDLG----- 1289
QY 331 NMTQFGHSSQTYDVTTPDTFVPHLGSIQANGISGNYIGVLSWV--SPFSHPGSGQVD 388
Db 1290 ----QVPLSSDNQ-----KTSSPDLGQVPLFPEDNQVFLDLQVPLSSDNQKTSSTD 1339
QY 389 LMK1-PNYG-----SSITEATHLAPSVYPGFGVLFVFFMSKIPGPA 430
Db 1340 LLTSLSPDFGQTVLSPDLQDLPLPSDNSQVTVSPDLSSLTSLSPDFNEIIL-----APDLGQ 1394
QY 431 YSL-PCLL----PQEYISHLAS-----EOA---PTVGEAALLHYVDPTGRTLGFEKAYPDG 479
Db 1395 VTLSPLDIQTNPALNHGHKASSADPDQASYPDPDSGQASSL-----PELNRTL-----PHPD- 1445
QY 480 FLTCTVP 485
Db 1446 -LTHIP 1450
```

Search completed: June 1, 2004, 13:57:54  
Job time : 14.6275 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 33.8809 Seconds  
(without alignments)  
4368.312 Million cell updates/sec

Title: US-09-926-799-2  
Perfect score: 2819  
Sequence: 1 MMASKATSSVDGASGAGQ.....YOLKPVGTASSARGRLGURR 530

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2793	99.1	530	14	US-10-314-739-3
2	244.5	8.7	668	14	US-10-209-507-4
3	238	8.4	669	14	US-10-209-507-2
4	237	8.4	547	12	US-10-670-695-4
5	237	8.4	671	12	US-10-670-695-2
6	126.5	4.5	1049	12	US-10-282-122A-49900
7	126	4.5	6310	12	US-10-282-122A-67793
8	124.5	4.4	913	15	US-10-369-493-2294
9	121.5	4.3	749	12	US-10-282-122A-55747
10	118.5	4.2	8360	14	US-10-132-134-34
11	116	4.1	2468	12	US-10-282-122A-66335
12	116	4.1	2468	14	US-10-246-330-4
13	112.5	4.0	878	14	US-10-157-031-171
14	112	4.0	702	12	US-10-087-192-366
15	112	4.0	769	12	US-10-424-599-212925

16	112	4.0	925	15	US-10-108-260A-3752	Sequence 3752, Ap
17	111.5	4.0	1765	12	US-10-282-122A-48055	Sequence 48055, A
18	111.5	4.0	6620	15	US-10-080-334-290	Sequence 290, App
19	111	3.9	1439	12	US-10-282-122A-44567	Sequence 44567, A
20	108.5	3.8	651	15	US-10-346-144-2	Sequence 2, Appli
21	108.5	3.8	769	14	US-10-029-386-33366	Sequence 33366, A
22	108	3.8	684	12	US-10-423-114-43059	Sequence 43059, A
23	108	3.8	724	15	US-10-346-144-5	Sequence 5, Appli
24	108	3.8	1465	15	US-10-369-493-13955	Sequence 13955, A
25	108	3.8	4952	15	US-10-051-874-56	Sequence 56, Appl
26	108	3.8	5159	15	US-10-085-198-112	Sequence 112, App
27	108	3.8	5262	15	US-10-051-874-165	Sequence 165, App
28	107.5	3.8	1048	14	US-10-283-538-10	Sequence 10, Appl
29	107	3.8	2358	12	US-10-282-122A-45763	Sequence 45763, A
30	106.5	3.8	668	12	US-10-282-122A-72916	Sequence 72916, A
31	106	3.8	792	15	US-10-369-493-4878	Sequence 4878, Ap
32	106	3.8	792	15	US-10-369-493-7637	Sequence 7637, Ap
33	106	3.8	3930	12	US-10-282-122A-46817	Sequence 46817, A
34	105.5	3.7	750	9	US-09-815-242-13985	Sequence 13985, A
35	105.5	3.7	750	12	US-10-282-122A-75580	Sequence 75580, A
36	105.5	3.7	752	12	US-10-282-122A-59788	Sequence 59788, A
37	105.5	3.7	996	14	US-10-101-464A-889	Sequence 889, App
38	105.5	3.7	996	14	US-10-101-464A-933	Sequence 933, App
39	105	3.7	407	14	US-10-156-761-12940	Sequence 12940, A
40	105	3.7	599	10	US-09-955-555A-29	Sequence 29, Appl
41	105	3.7	5262	15	US-10-051-874-167	Sequence 167, App
42	104.5	3.7	785	15	US-10-369-493-11883	Sequence 11883, A
43	104.5	3.7	5008	15	US-10-051-874-166	Sequence 166, App
44	103.5	3.7	353	10	US-09-946-374-397	Sequence 397, App
45	103.5	3.7	353	12	US-10-006-485A-397	Sequence 397, App

ALIGNMENTS

RESULT 1  
US-10-314-739-3

; Sequence 3, Application US/10314739  
; Publication No. US20030129588A1

; GENERAL INFORMATION:

; APPLICANT: ~~SECRET~~ ~~MAINTENANCE~~  
; Jiang, Xi

; Graham, David Y

; TITLE OF INVENTION: Methods and Reagents to Detect and

; Characterize No. US20030129588A1walk and Related Viruses

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Fulbright & Jaworski L.L.P.

; STREET: 801 Pennsylvania Ave., N.W.

; CITY: Washington, D.C.

; STATE: <Unknown>

; COUNTRY: USA

; ZIP: 20004

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/10/314,739

; FILING DATE: 09-Dec-2002

; CLASSIFICATION: <Unknown>

; PRIORITY INFORMATION:

; APPLICATION NUMBER: US/08/486,049

; FILING DATE: June 7, 1995

; ATTORNEY/AGENT INFORMATION:

; NAME: Davis, Peter

; REGISTRATION NUMBER: 36,119

; REFERENCE/DOCKET NUMBER: 311.023

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 202-662-0200

; TELEFAX: 202-662-4643

; TELEX: <Unknown>

```
; ORGANISM: Feline calicivirus
US-10-209-507-4
Query Match      8.7%; Score 244.5; DB 14; Length 668;
Best Local Similarity 32.2%; Pred. No. 1.6e-13; Indels 31; Gaps 8;
Matches 69; Conservative 27; Mismatches 87;

QY 19 GQLVPEVNASDPLAMPVAGSSTAVATAGQVNPIDPWIIINNFPVQAPQGEFTI-----SP 72
Db 136 GTLUVGGVIAEPNMQMSAVAD-----VATGKSVD-----EAFPSFTSVNMST 180
QY 73 NNTFGVGLFDLSLGPLNPFLLHLSQMYNGWGVNMRVIRIMLAGNAFTAGKIIVSCIPPGF 132
Db 181 SETQKILFKQSLGPLNPFLLHLSQMYNGWGVNMRVIRIMLAGNAFTAGKIIVVPPGI 240
QY 133 GSHNLITAAQLTLPFHVIADVRLDPIEVPLEDVNRVLFHNDRNQOTMELVCMLY----T 188
Db 241 DPVQST--SMLQYPHVLFDARQVEPVIFITPDLRNSLYHLS-DTDTTSLVIMIYNDLIN 297
QY 189 PLRTGGTGDSFVAVAGRVMTCPSPDFNFLVPP 222
Db 298 PYANDSNSGCIVT---VETKPGDFKFHLLKPP 328

RESULT 3
US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-2
Query Match      8.4%; Score 238; DB 14; Length 669;
Best Local Similarity 29.3%; Pred. No. 6.7e-13; Indels 38; Gaps 8;
Matches 67; Conservative 32; Mismatches 92;

QY 7 DATSSV---DGASGAGQLVPEVNASDPLAMPVAGSSTAVATAGQVNPIDPWIIINNFPVQA 63
Db 126 DGGSSITTEQGLVGGVIAEPSAQMTAADAATGKSD-----SEW----- 167
QY 64 PQGEFTI-----SPNNTPGGVLFDSLGLPHNPFLLHLSQMYNGWGVNMRVIRIMLAGNA 117
Db 168 -ESFSTFTSVNMSTSETQKILFKQSLGPLNPFLLHLSQMYNGWGVNMRVIRIMLAGNA 226
QY 118 FTAGKIIVSCIPPGFSGHNLITAAQLTLPFHVIADVRLDPIEVPLEDVNRVLFHNDRNQ 177
Db 227 VFGGKLAIAIVVPPGVDVQVST--SMLQYPHVLFDARQVEPVIFITPDLRSTLYHLS-DT 283
QY 178 QTMRLVCMLY----TPLRTGGTGDSFVAVAGRVMTCPSPDFNFLVPP 222
Db 284 DTTSLVIMIYNDLINPYANDSNSGCIVT---VETKPGDFKFHLLKPP 329

RESULT 4
US-10-670-695-4
```

```
; INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 530 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3
Query Match      99.1%; Score 2793; DB 14; Length 530;
Best Local Similarity 98.7%; Pred. No. 6.9e-252; Indels 0; Gaps 0;
Matches 523; Conservative 4; Mismatches 3;

QY 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMPVAGSSTAVATAGQVNPIDPWIIINN 60
Db 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMPVAGSSTAVATAGQVNPIDPWIIINN 60
QY 61 VOAPQGEFTISPNNTPGGVLFDSLGLPHNPFLLHLSQMYNGWGVNMRVIRIMLAGNAFTA 120
Db 61 VOAPQGEFTISPNNTPGGVLFDSLGLPHNPFLLHLSQMYNGWGVNMRVIRIMLAGNAFTA 120
QY 121 GKIIVSCIPPGFSGHNLITAAQLTLPFHVIADVRLDPIEVPLEDVNRVLFHNDRNQOTM 180
Db 121 GKIIVSCIPPGFSGHNLITAAQLTLPFHVIADVRLDPIEVPLEDVNRVLFHNDRNQOTM 180
QY 181 RLVCMLYTLRTGGTGDSFVAVAGRVMTCPSPDFNFLVPPTVQKTRPFTLPNPLSS 240
Db 181 RLVCMLYTLRTGGTGDSFVAVAGRVMTCPSPDFNFLVPPTVQKTRPFTLPNPLSS 240
QY 241 LNSRAPLPSMGISPDNVQSVQFONGRCTLDGLRGVTPVLSHVAKIRGTSNGTVIN 300
Db 241 LNSRAPLPSMGISPDNVQSVQFONGRCTLDGLRGVTPVLSHVAKIRGTSNGTVIN 300
QY 301 LTELDTGTPFPFPGAPIGPDLGGCDWHINMTQFGHSSQTYQVDVTTPTDTPFVPHLGSIQ 360
Db 301 LTELDTGTPFPFPGAPIGPDLGGCDWHINMTQFGHSSQTYQVDVTTPTDTPFVPHLGSIQ 360
QY 361 ANGISGNYIGVLSWSPSPHSGSQVDLWKINYPSSITEATHLAPSVYPPGFGVLF 420
Db 361 ANGISGNYIGVLSWSPSPHSGSQVDLWKINYPSSITEATHLAPSVYPPGFGVLF 420
QY 421 FMSKIPGAGYSLPCLLPQEVISHLASEQAPTVEAALLHYVDPDGTGRTIGEFKAYPDGF 480
Db 421 FMSKIPGAGYSLPCLLPQEVISHLASEQAPTVEAALLHYVDPDGTGRTIGEFKAYPDGF 480
QY 481 LTCVPNGASSGPQQLPINGVFVFSWVSRYQLKPVGTASSARGRLGLRR 530
Db 481 LTCVPNGASSGPQQLPINGVFVFSWVSRYQLKPVGTASSARGRLGLRR 530

RESULT 2
US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
```





QY	43	-----VATAGVNPIDPWIINNFIQAP-----	QG 66
Db	383	VCANANIITAAGSTITDNNGVLLSTGTSFTTTFDAGASTLGAITTGAGTTTTLHDDT	442
QY	67	EFTISP-----NTPGGVLEDLGLHLPFLHLS-OMYNGWGVNMVRIMLAGNAF	118
Db	443	DWTGNSATNLTSASNIHTAPADPSQAASYKTLOVQNTGADGNLHNLTLFGSDDA	502
QY	119	TAGKIVSCIPFGSHNLITIAQATLFPHVIAADVRLDPIEVLFDVNRVLFHNNDRNQ	178
Db	503	PSDKLIVSG-GRATGTNLTIHNTT-----GPGDMIVADGIQV-----	VQ 541
QY	179	TMLRVCLMLYPLRTGGT--GDSFVVAGVMTCPSPDNFL-----	217
Db	542	TL-----NGGTTASDAFTLTGLRL--AGAFDYLLFHGGIGGSRPNDFLRSS	586
QY	218	FLVPTVEOKTRPFTLNLPLSLNSRAP-----LPISGMGISPDNVQSVFQNGRCRTLD	273
Db	587	FVVPESI-----TPTVPLSPATSPSPGEGTSPSPG-EPGTSPSPGQCTPSSP	637
QY	274	GRLVGT-----TPVSLSHVAKIRGT-----SNGTVINLTBLD	305
Db	638	PGQGTPTSPPGQPGTSPSPGQPGTSPSPGQPGAPSPSPGQPGTSPSPGQGTTL	697
QY	306	GTTP-----HPPEGP-----APIGPDLGGCDWHINMT-----	333
Db	698	VSRFPFADPPASPLOQGVTVFVIGPELATYGVVQVVARQWFTMLGTLHERIGDILTADN	757
QY	334	-----QFGHSSQTOV-----DVTTPDTFVPHLGSIQANGIGSGNYIGVLSWVSPSPHPSG	384
Db	758	AGPDSGSHSGHARFQGQIDNRYQAFASPRATGQLLGVOAG-----	800
QY	385	SOVLMK-----IPN-----YGS-----ITEATHLAPSVYPGEGEVLPFFMSK	424
Db	801	--VDVNRGSLMPGHRDVAGVYFAYGNSLNDVDGLVTNATATA-----YLLGRGK	848
QY	425	IP-----GPGVSLPCLLPQEVYSHLASEQAPTVGEAALLHYVDVDPDTGRTLG	471
Db	849	LNLDAYSGGAYWTHYGGWGLDLVLQ-----TSVQGEATVQSARL-----PTTG	894
QY	472	EFKAYPQGLTCVPGNASSGQQPLPINGVVFVWSVSRFYQ-----LKPV--GTA	519
Db	895	-----DGFVTSLEAGY-----PVPPLGLFRFVLEPQIQVLMQVRSFQDENDGLGFDVLSGT	945
QY	520	SSARGRLGLR	529
Db	946	TGTTGRLGLR	955
RESULT 7			
US-10-282-122A-67793			
; Sequence 67793, Application US/10282122A			
; Publication No. US20040029129A1			
; GENERAL INFORMATION:			
; APPLICANT: Wang, Liangsu			
; APPLICANT: Zamudio, Carlos			
; APPLICANT: Malone, Cheryl			
; APPLICANT: Haselbeck, Robert			
; APPLICANT: Ohlsen, Kari			
; APPLICANT: Zyskind, Judith			
; APPLICANT: Wall, Daniel			
; APPLICANT: Trawick, John			
; APPLICANT: Carr, Grant			
; APPLICANT: Yamamoto, Robert			
; APPLICANT: Forsyth, R.			
; APPLICANT: Xu, H.			
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms			
; FILE REFERENCE: ELITRA.034A			
; CURRENT APPLICATION NUMBER: US/10/282,122A			
; CURRENT FILING DATE: 2003-02-20			
; PRIOR APPLICATION NUMBER: 60/191,078			
; PRIOR FILING DATE: 2000-03-21			
; PRIOR APPLICATION NUMBER: 60/206,848			
; Remaining Prior Application data removed - See File Wrapper or PALM.			
; NUMBER OF SEQ ID NOS: 78614			
; SOFTWARE: PatentIn version 3.1			
; SEQ ID NO 67793			
; LENGTH: 6310			
; TYPE: PRT			
; ORGANISM: Pseudomonas putida			
; US-10-282-122A-67793			
Query Match			
Beat Local Similarity 4.5%; Score 126; DB 12; Length 6310;			
Matches 122; Conservative 69; Mismatches 186; Indels 230; Gaps 31;			
QY	6	KDATSSVDGAS---GAGQL-----VPE-----VNASDPLA-----MDPVAGSST--AV	43
Db	3278	RDADGNVLGTATVGGDGRFDTLNAPORGESITVEATDNLGNSAGPVDFTAPDSPPQV	3337
QY	44	ATAGOVNPIDPWIINN-----FVQAPQGEFTISPNTGEGVLFDSLGL-PHNLPELH	95
Db	3338	VTNPAIDIGITITNGEPGATVTVRGDGVLSGTVVAGGA-FEITLDEPQTNGOALT	3396
QY	96	LSQMYNGWGVNMVR-----MLAGNAFTAGKIIVSCIPG-----FGS	134
Db	3397	VEQ--RDAAGNLSAAVLSAPDTPQAPDPTGLVLADG-----GSLSGSGEGFAGVTVGT	3450
QY	135	HNLITAOATLPHVIADVRLDPIEVLFDVNRVLFHNNDRNQTMRLVCLMLYPLRTGG	194
Db	3451	GGAVLGTATVQPDGTFQV-AIDPPQL-----NGQTLSTV-----QADNSGN	3490
QY	195	GTGDSFVVAGVMTCPSPDNFLFLVPPTEQKTRPFTLNLPLSLNSRAPLPSGMG	254
Db	3491	GSPAGTVIA-----PDF-----EAPLPAQGLG	3512
QY	255	ISPDNV-----QSQVQFQNGRCTLGRLVGTTPVLSL---HVAKIRGTSNGTVINLT	302
Db	3513	LDPTGTLTLLSGOGEAGSTVEVRNAA-----GNLLGTFPVNADGTFVALSPAQTNGEVL	3568
QY	303	ELDGTFFHPEGPAPIGFPDGLGGCDWHINMTQFGHSSQTOYVDVDTTPD-----TFVPHLG	357
Db	3569	LIDG-----GG-----NASPSATFTADDSTAPAAPAGLTITPGS	3603
QY	358	SIOANG-----IGSGNYIGVL-----SWVSP---PSHPSSGQVDLWKIPNYGSSIT	400
Db	3604	AIQGSGEAGSTVEVKLADGLTGLTVVVPAGGSFTVPLSPAQLDGOALNVLTDAAGNISQ	3663
QY	401	EATHLAPSVYPP-----GFGVLFVFFMSKIFGPGAY	431
Db	3664	PSQIFAPDITPPALPTDVAVSSDGTAVTGNAPGASSVTVSDGAGNVITVAVNP---DGSF	3720
QY	432	SLPCLLPQBYISHLASEQAPTV-----GEAALLHYVDVDPDTGRTLGEFKAYPDGFLTCV	484
Db	3721	SVPLDTPQN-----NGQTVTVVVTDAAGNDSAPSVTAPDT-----TNPEPATGLTVS	3768
QY	485	PNGASSG	491
Db	3769	PDGSTVG	3775

## RESULT 8

US-10-369-493-2294  
; Sequence 2294, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2294  
; LENGTH: 913  
; TYPE: PRT  
; ORGANISM: Schizosaccharomyces pombe  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(913)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-2294

Query Match 4.4%; Score 124.5; DB 15; Length 913;  
Best Local Similarity 19.6%; Pred. No. 0.044;  
Matches 111; Conservative 65; Mismatches 197; Indels 193; Gaps 26;  
QY 18 ACQLEPVNASPLAMDPAVAGSTAVAT-----AG--QVNPIDPWIINNFVQPGQEF 68  
Db 16 AGTVVNVQVEPEVEVTVPAVGTITSTVPEVVPXAGIVEVNGIDPETIVSVTVPAVGT 75  
QY 69 TISPNNTPGGVLFDSLGLPHLNPFLHLQSMYNGVGNMVRIMLAGNAFTAGKIIVSCI 128  
Db 76 T-----STVPDVVPX-----AGNVEVVP-----TEPDIMSVT 103  
QY 129 PGFGSHNLTIAQTLFPHV-----IADVTLDP-----IEVPLEVRNVLPHNDRNQ 178  
Db 104 VPAVGSTTSTV-----PDVEXAGIVEVNGIDPETIVSVTVPAVGA----- 144  
QY 179 TWRLVCMLYTPTLTGGTGGDSFVAGRVMTCPSPDENFLVPPVTEQKTRPTLPL 238  
Db 145 -----TISTVPDVVPXAGNVEVYPTPEPDIIIVSVTVPAVGSTTS--TVPDV-- 187  
QY 239 SSLSNSRAPLPSGM-----GISPDNVQSVQFQNGRCTLDGLVGTTP----- 281  
Db 188 -----VPXAGIVEVNGIDPETIVSV-----TVPAVATISTVDPVVLXAGNVEVVP 233  
QY 282 -----VSLSHVAKIRGTSNGTIVNLTELDG--TFPHPEGPAPIGFDPDLGGCDWHINMTQ 336  
Db 234 TEPDIMSVMVTVPAVGSTTSTVDPVVPXAGNVEVYPT-----PDI-----MVSVTVA 281  
QY 337 HSSQTYDVTTPDTFVPHLGSIQANGICSGNY-----IGVLSVSPSPSPSSQVDL 389  
Db 282 VGSTT-----STVPDV--VPXAGIVEVNGIDPETIVSVTVPAVGSTTMTAPDEVSAVIVDI 336  
QY 390 WK-----IPNYGSSITEATHLAP-----SVVP-----PGFGEVLVFFMS 423  
Db 337 KETDPEITVSVTVPAVGSTTSTVTVPLAGNVEVYPTPEPMVSVIVPAVGSTTSTVPE 396  
QY 424 KIPGGAISLPCLLPOEYISHL-----ASEQATVGBAALLHYVDPDTGRTLGEF 473  
Db 397 VVPLAGNVEVKVLEPETMVSIVPAVGSTTSTVPEVVPVPLAGNVE--VNALEPDITMSV--- 452  
QY 474 KAYPDGFLTCVNGASSGQQQLPING 499  
Db 453 -----IVPAVGSTTSTVPEVVPVLAG 472

## RESULT 9

US-10-282-122A-55747  
; Sequence 55747, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-03-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 55747  
; LENGTH: 749  
; TYPE: PRT  
; ORGANISM: Enterobacter cloacae  
US-10-282-122A-55747

Query Match 4.3%; Score 121.5; DB 12; Length 749;  
Best Local Similarity 21.3%; Pred. No. 0.062;  
Matches 120; Conservative 54; Mismatches 174; Indels 215; Gaps 33;  
QY 28 SDPLAMDVP-AGSSTAVTAGQVNPIDPWIINNFV---QAPQGETISGNNTPGGVLPDL 83  
Db 150 SDPKKITPVVRFSTVQGGAGSADTVRD--IRGFATKFTYEGIFDLVGNNTPVFFIQD- 206  
QY 84 SLGPHLNPFLHLQSMYNGVGNMVRIMLAGNAFTAGKIIVSCIPPFGSHNLTIAQAT 143  
Db 207 ---AHKFPDFVHAVKPEPHM-----AIPQGSADH-----T 234  
QY 144 LPPHVIADVTLDPIEVEPLEDRNVLPHNDRNQTMRLVCMLYTPTLTGGTG----- 197  
Db 235 FMDYV---SLQP-----ETLHNVWVWMSDRG-----IPRSYRTMEGFGIHTFRL 275  
QY 198 -----DSFV-----VAGRV-----MTCPSPDFN-----PLF 218  
Db 276 INAEKATVRFHWKPVAGKASLVWDEAQKLTGRDPDFHRELWESIEAGDPPELVGLQ 335  
QY 219 LVPPTVEQKTRPFTLPNL-PLSSLSNSRAPLPSG---MGISPDNV-----QSVQFQNGRC 270

314 ---GPAPIGPDPOLGGDMHNMTOF--GHSSQTVQYVDVTPDTFVPHLGSIOANG-----I 364  
1714 RVGDPAGSVVLGSKVTNLGHTQFAAGIAGVLKALLALQEQRVPPSLHFAEANARVPLD 1773  
365 GSNYIGVLG--WVSPSPSHPSQVDLWIKIPNYGSSITEATHLAPSVYPP-----412  
1774 GSPFTVATTAAQMPPEPAEGFRRAVS-----SFGASGTNA-HVLEEHPPVVRATTGPESA 1827  
413 -GGEVLVFFMSKIPG-----PG-----AYSLPC-----435  
1828 GGDGEAAFLSLATPAALRAVAERLLAREREGLPARQAVAYSLAAGRHRFPHRLAVVAT 1887  
436 ---LLPOEYISHLASEQAPTVGSAALLHYVDPTGR--TLGEF-----KAYPDG-----F 480  
1888 GLPALAARLARWLADEQPG--GEGTLLHGVAHAGTRQAALGGLAPAEALAAAYVGGAGPPF 1945  
481 LTCVPNGASSGPQQLPI 497  
1946 AESFPAGAR---RQVPL 1959  
RESULT 11  
US-10-282-122A-66335  
; Sequence 66335, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 66335  
; LENGTH: 2468  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-10-282-122A-66335

336 LIPEDEFEK---FDFDLDDPTKLIPEELVPQVLGKMLNRPDNFFAENEQVAFHPGHI 392  
271 T-----LDGLRGVTTVPVSLGHVAKIRGTSNGTVINLTDLDTGTFPHFPEGAPITGF 320  
393 VPGLOFTNDPLQGLRFSYDTQIS-----RLGGPNFHEI-----PINR 431  
321 POLGGCDMHNMTOFCHSSQTVQYVDVTPDTFVPHLGSIOANGISGNYIGVLG--WSP 378  
432 PT---CPYH---NFQDGMHRQDIDTNPANYPN--SINDN-----WPRETP 470  
379 PSHPSG-----SQVDLWIKIPNYGSSITEATHLAPSVVPPGFEVL-----VFMSKIP-- 426  
471 PGPKRGGFESYQERID-----GNKIRERS-----PSGEYVAHPLRFWNSQTFIE 515  
427 -----GPGAYSLPCLLPQSYISHLASEQAPTVGSAALLHYVDPTGRTLGE-----472  
516 QQHIIIGGFSELSKVY-RYIRERVVDQLAHI-DVQLAQSVADNLGITLTDQRNLAPPK 573  
473 ----FXAYPDGFLTVCVPGASSG 491  
574 EVNGVKKDPSSLVYAVPGGSIK 596  
RESULT 10  
US-10-132-134-34  
; Sequence 34, Application US/10132134  
; Publication No. US20030171562A1  
; GENERAL INFORMATION:  
; APPLICANT: Farnet, Chris  
; APPLICANT: Yang, Xianshu  
; APPLICANT: Staiffa, Alfredo  
; APPLICANT: Zazopoulos, Emmanuel  
; TITLE OF INVENTION: POLYKETIDE SYNTHASE ENZYMES  
; FILE REFERENCE: 3012-2US  
; CURRENT APPLICATION NUMBER: US/10/132,134  
; CURRENT FILING DATE: 2002-04-26  
; NUMBER OF SEQ ID NOS: 43  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 34  
; LENGTH: 8360  
; TYPE: PRT  
; ORGANISM: Streptomyces amphibiosporus  
US-10-132-134-34  
Query Match 4.2%; Score 118.5; DB 14; Length 8360;  
Best Local Similarity 23.2%; Pred. No. 4.8;  
Matches 143; Conservative 56; Mismatches 241; Indels 177; Gaps 36;  
QY 13 DGASGAGQLVPEV-----NASDPLAM-----DPVAGSSTAVATAGQVNPIDPWI 56  
DB 1388 DGTGAARFVDPADVPAPAREADDPVAVVGMAARYGDAEDPRALMDRLLAGDLDLVEPVRWD 1447  
QY 57 INNFOVQAPQGEFT-----ISPNNTPGGVFLDLS--LGPHLNP-FLHLSSQYN-----GW 103  
DB 1448 LGPEVTCRAGSFVRGMDRFD-----VFFAISGVEAAHMDPQRIEFLQCNWNALEDAGY 1501  
QY 104 VGNMRVIRMLAGNAFTAGKIIVSCIPPGFGSHNLT-----AQATLPHVIAVRLDPI 158  
DB 1502 TGE-RLRERNCG-----VYVGCVY-----GDYDSIGDRAPAQA-LWGTGSGVVASRIY 1549  
QY 159 EVPLEDVNRNVLFHNNDRNQOTMRLVCLMYTLPTRTGGGTGDSFVVAGRVMTCPSPDNFLF 218  
DB 1550 QDLKGPALTDTTSCSSLSVSLHLAC-----RLRTGAA--DWAJAGGV-----FLQ 1594  
QY 219 LVPPTVEQKTRFTPLNPLSLSNSRAP--LPTSGMI-----SPDNVQSVQF 265  
DB 1595 TTPRLYEATRAGMLSPTRCHSFDSRAGDFVPGEGAGVVLKRLSDALRDGDHVGVLVR 1654  
QY 266 QNGRCITLQRLVG-TTPVSLSHVAKIRGTSNGTV---INLTDELGTTPH---PFE-----313  
DB 1655 ATG-VNQDGTNGITAPSAQSQALLREVHAGVPGVQLVEAHGTGTQLGDPIEFRLS 1713



Wed Jun 2 09:13:26 2004

us-09-926-799-2.rapb

```
Db 232 PSTIWSSTPTIIMSSSPSSASITVPFATIHSPSSPYIFSTENVGSASITAFPSLSS 291
Qy 399 ITEATHLAPSVYPPGFGVLPV-----FMSKIPGCGAYSLPCLLPQEIYISHLASEQAPTVG 454
Db 292 STTSTPSSSTTALTITPESYISLSPSTTPCPGTITI-TIVP-----ASTPDCV- 342
Qy 455 EAALLHYVDPDGT-----LGEFKAYPDGFLT 482
Db 343 -----EMDPSTEATSPPTPLTVFPFITEMVTCPSISMQTTLATHMDTSSMTPESSS 396
Qy 483 CVPNGASS-GPOOLPINGVFPV-----SWVS 508
Db 397 IIPNASSSTGTGTVPNTVFTSTRLPTSETWLS 429

RESULT 14
US-10-087-192-366
; Sequence 366, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 366
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-087-192-366

Query Match 4.0%; Score 112; DB 12; Length 702;
Best Local Similarity 23.7%; Pred. No. 0.43;
Matches 75; Conservative 41; Mismatches 114; Indels 86; Gaps 16;

Qy 210 PPDFNFLEL-VPTVQKTR-PTLPNPLSSLSNRAPLPISGMGISPDNVQSVQFQN 267
Db 290 PSNGVEFISVQSPPTVSLTKNPVSLFSLPNPTKPNKVPSPSPSIQRP- 339
Qy 268 GRCTLDGLVCTT-----PVSLSHVAKIRGTSNGTVINLTDLGTFPHFEGPAPIGFP 321
Db 340 ---TASAPLGTTLAVQAVPTAHIVQATR-TSLPTV-----GPSGLYSP 380
Qy 322 --DLGGCDWHINMTQGHSSQTVQDVTPTTFVPHLGSIQANGIGSGNVIG-VLSVW-- 376
Db 381 STNRGPIQMKIPISAFSTSSAAEQNSNTTPRIDPKLNHTPVSTMSSQFVSRPLQIQP 440
Qy 377 SPPHPSC-----SQVDLWKIPNYGSS-----ITEATHLAPSVYPPGFGVLPVFFMS 423
Db 441 APPLQPSGVPTSGPSQTTIHLPTAPTIVNTHRPVQTVTRLPVPAPANHQV-----YT 497
Qy 424 KIPGPGAYSLPCLLPQEIYISHLASEQAP---TVGEAALLHYVDPDGTGRTLGEFKAYPDGF 480
Db 498 TLPAFPA-----QAPLRGTVMQAPAVRQVNPQNSVTV-----RVPQTT 535
Qy 481 LTCVPNG-----ASSGPO 493
Db 536 TYVNNGLTLGSGTGPQ 551
```

RESULT 15  
US-10-424-599-212925  
; Sequence 212925, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:

```
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 212925
; LENGTH: 769
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(769)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_34297C.1.pap
US-10-424-599-212925
```

```
Query Match 4.0%; Score 112; DB 12; Length 769;
Best Local Similarity 21.3%; Pred. No. 0.5;
Matches 73; Conservative 43; Mismatches 116; Indels 110; Gaps 17;

Qy 232 TLPNPLSSLSNRAPLP-SCMGISPDNVQSVQFQNGRCRTLDGRLVGTTPVSLSHVA 288
Db 356 TFSEMPSSSQTCAYPAHTAFSNGMGLSLNL-----TMDG-----METDETR 398
Qy 289 KIRGTSNGTVINLTDLGT-----PFHFGPAPIGFPDLGGCDWHI 330
Db 399 PAEGSQHGLNTNTYSLNGLMLHLSRQTANRGVLSSEFQGFQHP-----PFSRDPGWEI 451
Qy 331 NMTQFGHSSQTVQDVTPTTFVPHLGSIQANGIGSGNVIGLSWVSPSSH----- 381
Db 452 PFLHGWINGQSVQGV-----PSMLPHMGASRDN---LSQHIGSSSYIKASNHSTSNVDAAM 504
Qy 382 -----PSGSQVDLWK-----IPNYGSS-----ITEATHLAPSVYPPGFGVLPVFFMSKIPGP 428
Db 505 PSSAIPGIGISQTI SMRSGLENHFSQSRVPVSESGNLAASINTPHDGFDMQTIIVSRIOSE 564
Qy 429 GAYS-----LPCLLPQEIYISH-LASEQAPTVGEAALLHYVDPDGTGRTLGEFKAYPDGF 480
Db 565 LATSVAATAAELPCTVKLRVMSHDIKNPCAP-----LNAADP-----GL 603
Qy 481 LTCVPN--GASSGPOOLPINGVFPVSVWSRFFYQLKPVGTAS 520
Db 604 QTPVHQDPGVATSPTRHFI SAHQV---WDAR--MLPPLEEAT 640
```

Search completed: June 1, 2004, 14:04:29  
Job time : 36.8809 secs



Db 421 FMSKMPGAYNLPCLLPQYIYSHLSEQAPTVEAALLHYVDPDTGRNLGEFKAYPDGF 480  
 QY 481 LTCVNGASSGQQLPQYIYSHLSEQAPTVEAALLHYVDPDTGRNLGEFKAYPDGF 530  
 Db 481 LTCVNGASSGQQLPQYIYSHLSEQAPTVEAALLHYVDPDTGRNLGEFKAYPDGF 530

RESULT 2  
 B37491  
 major capsid protein [similarity] - Southampton virus  
 N/Alternate names: orf2 protein  
 C/Species: Southampton virus  
 C/Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 18-Aug-2000  
 C/Accession: B37491  
 R/Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
 Science 259, 516-519, 1993  
 A/Title: Sequence and genome organization of a human small round-structured (Norwalk-like) virus  
 A/Reference number: A37491; MUID:93142023; PMID:8380940  
 A/Accession: B37491  
 A/Status: not compared with conceptual translation  
 A/Molecule type: nucleic acid  
 A/Residues: 1-546 <LAM>  
 A/Cross-references: GB:07418; NID:gl236787; PIDN:AAA92984.1; PID:g295114  
 A/Note: sequence extracted from NCBI backbone (NCBI:123458)  
 C/Species: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype  
 C/Superfamily: human calicivirus capsid protein  
 C/Keywords: Glycoprotein  
 F/303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 72.1%; Score 2033.5; DB 2; Length 546;  
 Best Local Similarity 69.6%; Pred. No. 3.5e-136; Mismatches 82; Indels 17; Gaps 6;  
 Matches 380; Conservative 67;  
 QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPVAGSSTAVATAGQVNPIDPWNNF 60  
 Db 1 MMASKDAPQSDAGASGAGQLVPEVNTADPLPWEVAGPTTAVATAGQVNMIDPWNNF 60  
 QY 61 VQAPQGEFTTSPNTPGGVLFDSLGLPHLNPFLHLSQMYNGVGNVRVIMLAGNAFTA 120  
 Db 61 VQAPQGEFTTSPNTPGGVLFDSLGLPHLNPFLHLSQMYNGVGNVRVIMLAGNAFTA 120  
 QY 121 GKIVSCIPPGFGSHNLTIAQATLPHVIAVDRVLTDPVLEPVDVRLVFNHNDNQOTM 180  
 Db 121 GKIVCCVPPGFTSSSLTIAQATLPHVIAVDRVLTDPVLEPVDVRLVFNHNDNQOTM 179  
 QY 181 RLVCMLYPLRTGGGTG--DSFVAVAGVMTCPSPDFNPLFLVPTVEQKTRPFTLPL 238  
 Db 180 RLVCMLYPLRTGGGSGNSDSFVAVAGVMTCPSPDFNPLFLVPTVEQKTRPFTLPL 239  
 QY 239 SSLNSRAPLPISGMISPDNVQSVQFONGRCTLDGRLVGTTPVSLSHVAKIRGTSN--G 296  
 Db 240 QTLNSRFPFSLIQGMILSPDASQVQVQNGRCLDGLGTLTPATSGQLFRVGRKINQGA 299  
 QY 297 TVINLTDLGTPFHFEGPAPIGFDPDLGGCDWHNMTQFGHSS-----QTYQYVDTTPTD 351  
 Db 300 RTLNLTVDGKRFMAFDSPAPVGFDFKCDWHMRISKTNNKSGDPMRSVSVQTNVQ 359  
 QY 352 FVPHLGSIQANGI---GSGNYIGVLSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 408  
 Db 360 FVPHLGSIQFDEVNHPTGDIYGTIEWISQSPSTPGTIDINLWEIPDYGSSLSQAANLAPP 419  
 QY 409 VYPPGEGVLVFFMSKIPGPGAYS---LPCLLPQYIYSHLSEQAPTVEAALLHYVDP 464  
 Db 420 VFPFGEGEALVYFSAFPFGNNRSAPNDVPCLLPQYIYTHFVSEQAPTMGDAALLHYVDP 479  
 QY 465 DTGRTLGEFKAYPDGFLTCVNGASSGQQLPQYIYSHLSEQAPTVEAALLHYVDP 524  
 Db 480 DTRNLGEFKAYPDGFLTCVNGASSGQQLPQYIYSHLSEQAPTVEAALLHYVDP 539  
 QY 525 RLGLRR 530  
 Db 540 RLGVRR 545

RESULT 3  
 S60616  
 capsid protein - human calicivirus (strain Melksham)  
 C/Species: human calicivirus  
 A/Variety: strain Melksham  
 C/Date: 23-May-1997 #sequence\_revision 23-May-1997 #text\_change 28-Jul-2000  
 C/Accession: S60616  
 R/Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
 Virus Res. 37, 271-283, 1995  
 A/Title: Capsid diversity in small round-structured viruses: molecular characterization  
 A/Reference number: S60615; MUID:96136658; PMID:8533462  
 A/Accession: S60616  
 A/Molecule type: genomic RNA  
 A/Residues: 1-542 <GRE>  
 A/Cross-references: EMBL:X01879; NID:g976077; PIDN:CAAS7462.1; PID:g976079  
 A/Note: it is uncertain whether Met-1 or Met-3 is the initiator  
 C/Superfamily: human calicivirus capsid protein  
 C/Keywords: capsid protein; coat protein  
 Query Match 43.9%; Score 1238; DB 2; Length 542;  
 Best Local Similarity 47.5%; Pred. No. 6.6e-80;  
 Matches 264; Conservative 74; Mismatches 168; Indels 50; Gaps 13;  
 QY 1 MMASKDATSSVDGASGAGQLVPEVNASDPLAMDVPVAGSSTAVATAGQVNPIDPWNNF 60  
 Db 1 MMASNDAAAPSTDCGAG--LVPESN-NEVMALEPVAGAAAPVGTQNTIIDPWNNF 56  
 QY 61 VQAPQGEFTTSPNTPGGVLFDSLGLPHLNPFLHLSQMYNGVGNVRVIMLAGNAFTA 120  
 Db 57 VQAPNGEFTTSPNAPGEVLLNLELGPENLAPYLAHLMYNGVAGNEVQVMLAGNAFTA 116  
 QY 121 GKIVSCIPPGFGSHNLTIAQATLPHVIAVDRVLTDPVLEPVDVRLVFNHNDNQOTM 180  
 Db 117 GKLVFAVPPHPPVENLSPOKIMTFPHVIVDRTLEPVLPLPDRVNSFFHYNQDDPKM 176  
 QY 181 RLVCMLYPLRTGGGTGDSFVAVAGVMTCPSPDFNPLFLVPTVEQKTRPFTLPL 240  
 Db 177 RIVAMLYPLRSNGSGDDVFTVSCRVLTRPSPDFLYLVPPTVESKTKPTLFTLIGE 236  
 QY 241 LNSRAPLPISGMISPDNVQSVQFONGRCTLDGRLVGTTPVSLSHVAKIRGTSN-- 298  
 Db 237 LNSRFPFPIQGMISPDNVQSVQFONGRCTLDGRLVGTTPVSLSHVAKIRGTSN-- 296  
 QY 299 -----INLTDLGTPFHFEGPAPIGFDPDLGGCDWHI-----NMTQFGHSSQTYQYD 346  
 Db 297 NDHLNNVTITNLGSGDFPSEIDIPALGVDPFQGRVFGVISQRDKQNAAGHSEFANRGHD 356  
 QY 347 TTPDT-----FVPHLGSIQANGISGNYIGVLSW-----VSPPSH--PSG---SQVDLW 390  
 Db 357 AVVPTVTAQYTPKLGQIQ-----IG--TWQTDLTIVNQPVKFTPVGLNDTEHFNQW 405  
 QY 391 KIPNYGSSITEATHLAPSVPVPGFGEVLFVFFMSKIPGPGAYSPLP---CLLPQYIYSHLAS 447  
 Db 406 VVRYAGALNLTNLTNAPSAPVFPGERLLFFRSHLPLKGGYGNPAIDCLLPQEWQHFYQ 465  
 QY 448 EQAPTVEAALLHYVDPDTGRTLGEFKAYPDGFLTCVNGASSGQQLPQYIYSHLAS 507  
 Db 466 EAAPSMSEVALVRYINPDTRALFEAKLHRAFGMTVSSN--TSAPVVPVFPANGYFRFDSWV 523  
 QY 508 SREYQKPKVGTASSAR 523  
 Db 524 NQFYSAPMGTNGRR 539  
 RESULT 4  
 S40111  
 capsid protein - human calicivirus (strain Bristol isolate B493)  
 C/Species: human calicivirus  
 A/Variety: strain Bristol isolate B493  
 C/Date: 25-Dec-1994 #sequence\_revision 27-Feb-1997 #text\_change 28-Jul-2000  
 C/Accession: S40111



R;Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
submitted to the EMBL Data Library, December 1993  
A;Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de  
A;Reference number: S40111  
A;Accession: S40111  
A;Molecule type: Genomic RNA  
A;Residues: 1-539 <GRE>  
A;Cross-references: EMBL:X76716; MID:G436410; PIDN:CAA54134.1; PID:G436411  
A;Experimental source: human enteric calicivirus strain Bristol isolate B493  
C;Superfamily: human calicivirus capsid protein  
C;Keywords: capsid protein; coat protein

Query Match	40.7%;	Score 1148.5;	DB 2;	Length 539;
Best Local Similarity	45.1%;	Pred. No. 1.4e-73;		
Matches	249;	Conservative	82;	Mismatches 182;
			Indels	39;
			Gaps	14;
QY	1	MMWASKDAYTSSVDGASGAGOLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWITNNF	60	
DB	1	MMWASNDANPS-DGS--AANLVPEVN-NEWMALEPVVGAIAAPVAGQONVIDPWIRNRF	56	
QY	61	VQAPQGEFTISPNNTFGGVLFDLSLGHLPNPFLLHLHSQMYNGVGNMVRIMLAGNAFTA	120	
DB	57	VQAPGGEFTVSPRNAPGEILWSAPLGPDLNPNVLSHLSRMVNGYAGGFEVQVILAGNAFTA	116	
QY	121	GKLIIVSICIPFGSGHNLTIAQATLPHVITADVRLTDPIEVLPELDVRYNVLPHNDRNQOTM	180	
DB	117	GKVFIAAIVPNPFTEGLSPSQVMTMFPHIIVDRVQLSEPVLIPLPDVRNNFYHYNQANDSTL	176	
QY	181	RLVCMLYTPLRITCGGTGDSFVAVGRVMTCPSPDFNFLVLPVPTVEOKTRPFTLPNLPSS	240	
DB	177	KLIAMLYTPLRANNAGDDVFTVSCRVLTPSPDPDFILVLPVPTVESRTKFTFVPLTVBEE	236	
QY	241	LSNSRAPLPISGMGIISPDNVQSVQFONGRCTLDGRLVGTTPVLSLSHVAKIRGT-----S	294	
DB	237	MSNSRFPPIPLEKLYTGPFSAFVVPQNGRCTTDGVLGTTQLSNAVNICNFRGDVTHIAGS	296	
QY	295	NGTVINLTLDGTPPEPFEG-PAPIGFDPDLGGCDWHNNMTQF-----GHSSQTQYDVTTP	349	
DB	297	HDYTMNLASQNSNYDPTBEIPAPLGTDPFGKIQGL-LTQTTRADGSTRAHKATVSTGS	355	
QY	350	DTFVPHLGSITQ-----ANGTSGNY-----IGVLSWVSPFSPHSPGSQVDLWKPINYGSSI	399	
DB	356	VHFTPKLGSVQFTDTNDNFQAGQNTKFTPVGVIQ--DGDHHQNEPQ--QWLLPNVSGRT	411	
QY	400	TEATHLAPSVYPGCGEVLVFFMSKIPGCGAY---SLPCLLPQEIYISHLASQAPVTGBA	456	
DB	412	GHNVHLAPAVAPTTPGGEQLLFFRSTMPGCGSGYPNNMLDCLLPQEWVHLHFYQEAAPQSDV	471	
QY	457	ALLHYVDPDGTGRLTGFFKAYPDGFLTCVPNGASSGPOQ--LPIINGVFVSVMSVRFYQLK	514	
DB	472	ALLRFVNDPTGRLVFPCKLHKSQYITV-----AHTGPYDLVLPNGVFRFDSWVNGFYTLA	527	
QY	515	PVGTASSARGRL	526	
DB	528	PMNGTGRRRAL	539	

RESULT 5  
A53982  
capsid protein - European brown hare syndrome virus  
C:Species: European brown hare syndrome virus  
C>Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 28-Jul-2000  
C/Accession: A53982  
R:Wirllich, C.; Meyers, G.; Ohlinger, V.F.; Capucci, L.; Eskens, U.; Haas, B.; Thiel, H.  
J. Virol. 68, 5164-5173, 1994  
A;Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease  
A;Reference number: A53982; MUID:94309183; PMID:7518531  
A/Accession: A53982  
A;Status: preliminary  
A;Molecule type: mRNA  
A;Residues: 1-576 <WIR>  
A;Cross-references: GB:U09199  
C;Superfamily: human calicivirus capsid protein

Query Match	13.0%	Score	366.5	DB 2	Length	576			
Best Local Similarity	26.0%	Pred. No.	3.3e-18						
Matches	137	Conservative	81	Mismatches	221	Indels	87	Gaps	21
QY	9	TSSVDCGASGAGQLVPEVNASDPLAMPDVAGSSATAVAG-----QVNPDPDWIINNFPVQ	62						
DB	15	TASVPGTTDGMDPGVVASTDVTADNVAAS-----VATAGIGGGPQQASQESWRVNFYF-	70						
QY	63	APOGBETISPNNTPGGVLFDLSLGPHLNPFLLHLSQMYNGVMGNMVRIMLAGNAPTAGK	122						
DB	71	--NDVFTWSVTDAPGSGILYSVQHS PQNNPFTQLV SQMYACWAGMQPFFIVAGSGIFGGR	128						
QY	123	IIVSCITPPGEG-SHNLTIAQAATLPPHVIADVRTLDPIEVLDPEDVRNVLFHNDRNRQQTMR	181						
DB	129	LVCAIITPPGIQIQGLEVRQ---FPHVIVDARSLEPVITITMPLDRPMEHYPTGDPGLVPT	185						
QY	182	LVCMLYTPI-RTGGGTGDSFVAGRVYMTCPSPDNFLFLVPPTVE--OKTRPPTLPNLP	238						
DB	186	LVSVYNNLINPEGGTISAIQVT--VETRESEDFEVLIRAPSKTVDSVNPSPWLLTTPV	243						
QY	239	--SSLSNSRAPLPISGMGISPDNVQSVQFQNGRCTLDGLRVGTTPVSL-----SHVAK	289						
DB	244	LTGAGSDNRWGAPIVGLQVPVGGFST--SNRHNNMGSTYGVSSPFDDIDHPSGNVSY	300						
QY	290	IRGTSNGTWINLTDLGT-PFHPPEGPAPIGPPDLGCDWH-----INMTQFGHSS	339						
DB	301	PTGSATNTIETVANAGTATTNPISINAPDGFDPDMAIPESGTTIPTGAWVGFQGVWNAS	360						
QY	340	QTOYDVTTPDTVPVPHLGSTQANGISGNVIGVLSWVSPSHSPSGSOVDLWKIPNYGSSI	399						
DB	361	NG-----TPYGVTVQAVELFAN--GAPSSIRPVTTTGAQLVAKSI--YGVAI	405						
QY	400	TEATHLAPSVYPGFGEVLPFFMSKIPGPGAYSLPCLL-----POEYIHLASEQAPT	452						
DB	406	AQNQSAG-----IIFLSK---GMVSTPGVAATYTPQPSAIVTTPCTPVAAP	450						
QY	453	VGEAALLHYVDDPTGRTLGEFKAYPDGFLTCVPNGASSGPGQQLPIN	498						
DB	451	IGNTPTIME--SAVVARTGDVNAAGGSV-----NGTOYGVGSOPLS	489						

```

RESULT 6
S55399
genome polyprotein - rabbit hemorrhagic disease virus (isolate BS89)
C:Species: rabbit hemorrhagic disease virus
A:Variety: isolate BS89
C:date: 15-Jul-1995 #sequence_revision 01-Sep-1995 #text_change 18-Jun-1999
C:Accession: S55399
R:Rossi, C.
submitted to the EMBL Data Library, May 1995
A:Reference number: S55399
A:Accession: S55399
A:Molecule type: genomic RNA
A:Residues: 1-2344 <ROS>
A:Cross-references: EMBL:X87607; NID:G9854640; PID:G9854641
A:Experimental source: isolate BS89
C:Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C:Keywords: polypeptin

```

	Query Match	11.6%	Score 326;	DB 2;	Length 2344;	
	Best Local Similarity	25.7%;	Pred. NO. 1.6e-14;			
	Matches 139;	Conservative 66;	Mismatches 171;	Indels 164;	Gaps 27;	
Qy	4	ASKDATSSVDGASGACQLVPEVNASDPLAMPVAGSSTAVATAG-----QVNPIDPWII	57			
		: :	:	:	:	:
		: :	:	:	:	:
Db	1779	AGTATTASVPGTTDG-LDPGVVATTSV--TAENSASIIATAGIGPPQOVQOETWRT	1835			
		:	:	:	:	:
Qy	58	NNFVQAQCEFTISPNNTGGVLFDLSGLPHLNPELLHSOMYGVGNMRVRIMLAGNA	117			
		:	:	:	:	:
Db	1836	NFY----NDVFTWSVADAPGSILITYQHSPQNPFPTAVLSQMTAGVAGGWQFRFIVAGSG	1892			
		:	:	:	:	:
Qy	118	F7AGKIIIVSCIPPGFG-SHNLITIAQTALTPSHVIADVRTLDPIEWPVEDRVNVLFIHNDRN	176			
		:	:	:	:	:



Db 1883 QRRFIVAGSGVGLRAVIRPPIGIEIGLEVRQ---FPHVVIDARSLEPVITITPDLR 1939  
QY 167 NVLFHNDRNQOQTMRLVCMLYTLRTGGTGDGDSFVVGVRVMTCPSPDFNLFPLVPP---T 223  
Db 1940 PNMVHTGDPGLVPTLVLSVNNLINPFG-GSTSAIQVTVETRPSEDFEFVIRAPSSKT 1998  
QY 224 VEQKTRPFTLPLNPLSSLSN-SRAPLPISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPV 282  
Db 1999 VDSISPAGLTTPVLTVGNGDRNRNGQIVGLQVPFGSTC---NRHMNLNGSTYGMSSP 2055  
QY 283 SLSHVAKIRGTGNTVINTEL-----DGTTP---FHPFEGPAPI 318  
Db 2056 RGDDIDRRGASYSYSGSNATNVLQFWYANAGSAIDNPISQVAPDGFPMDFPFGP--- 2112  
QY 319 GPPDLG----GCDWHINMTQFGHSSQTYQYDVTTPDFVPHLGSIQANGISGNGYIGVLS 374  
Db 2113 GIPAGAGWVGAIWNSN-----SGAPNVTTVQAVELGPAT--GAPG 2151  
QY 375 WYSPSPHSGSQVDLWKIPNYGSSITEATHLAPSVY-----PPGGEVLVFPMSK- 424  
Db 2152 NLQPTTNTSQAQT-----VAKSIYAVVTGTAQNPAQ-----LFVMSG 2189  
QY 425 -IPGCGAYSLPCL-LPQEVISHLASQAPTGEAALLHYVD-----PDTGRTLGRFKAYP 477  
Db 2190 IISTPNAQRITVTPQDDRIVTTPTGTPAAPVGNTPIMPASVVRTGDVNNATAGS----- 2244  
QY 478 DGLFCLCPNGA--SSGPQOLPI 497  
Db 2245 -----ANGTQGTGSQLPV 2259

## RESULT 9

C48562  
coat protein - San Miguel sea lion virus (serotype 1)  
N:Alternate names: capsid protein  
C:Species: San Miguel sea lion virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
C:Accession: A48562  
R:Neill, J.D.  
Virus Res. 24, 211-222, 1992  
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus  
A:Reference number: A48562; MUID:92410750; PMID:1529644  
A:Accession: A48562  
A:Molecule type: genomic RNA  
A:Residues: 1-702 <NEI>  
A:Cross-references: GB:M87481; NID:G334882; PIDN:AAA16217.1; PID:G334884  
A>Note: sequence extracted from NCBI backbone (NCBIN:113564, NCBI:113565)  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 9.3%; Score 261.5; DB 1; Length 702;  
Best Local Similarity 23.9%; Pred. No. 1.2e-10;  
Matches 137; Conservative 59; Mismatches 222; Indels 155; Gaps 26;  
QY 13 DGAGAGOLVPS---VNASDPLAMPVAGSSAVATAGQVNPIDPWIINNFFVOAPGGEFT 69  
Db 154 DPGGADIIVTEQGTVVOQPV---PAQSALTTLAAASTGKTVD-----CEWT 198  
QY 70 I-----SPNTPGGLVFDLGLPHLNPFLHLSQMYNGVGNMVRIMLAGNAPT 119  
Db 199 TFFSYHTAVNWSSTEAQKILPSRALSPELNPLRHISLSYTWSGGIDVFTVSGGVF 258  
QY 120 AGKIIIVSCIPPGFGS-HNLTIAQATLPHVIADVRLDPIEVLVDVRLVFNHNDNRQO 178  
Db 259 GCKLAALIVPPGIEPVESPTMLQ---YPHVLFDARQTEVFIFIPDIRKTLYHSD-DTD 314  
QY 179 TMLVCMLYTLP-----RTGGGTGDSFVVGVRVMTCPSPDFNLFPLVPPVTEQKTRPFTL 233  
Db 315 TTRLVIMVNLINPYEQSEPKSSCSITVETR-----PSSDFTFLSKKPGSLKHGSLPS 370  
QY 234 PNLPLSS--LSNSRAPLPISGMGISPDNVQSVQFQNGRCTLDGRLVG-TTP-----VSL 284

Db 371 DLIPNRSRHMGNRRWSTIDGFVQPRVFQS-----NRHFDFFDSTTTGWSPTYIPIEVTL 426  
QY 285 SHV-----AKIRGTSNGTVINLTDELDTGTFHPFEGPAPIGPPDL-----GGCDW 328  
Db 427 EKLDRGGQYFKYTDTEKSLVPCLE-----PDGWPDTTIPMTAMTASNGYDY 471  
QY 329 HI-----NMTPQ-----GHSSQT-----QYDVTTP 349  
Db 472 TVAEYRIITNNGTHFGFYIMGNLTTKVAGSDNLGETQTSRTLFASVGNKQDQNTINPTH 531  
QY 350 DTFVPHLGSIQANGISGNGYIGVLSWSPSPHSPSSQVDLWKIPNYGSSITEATHLA--P 407  
Db 532 KITSNLVVYDANNVSAAT-AKTITWSTMTSHGLVYLVDSESV---GSDSTKVRIATLUP 587  
QY 408 SVYPPGFGFVLFVFMFKIPGPGAYSLPCLLPQEVISHLASQAPTGEAALLHYVDPDG 467  
Db 588 EAFYNG-GNFPVFFTNKI-----QIGHFDRAHTKCFNSQVLM-----TS 625  
QY 468 RTLGE--FKAYPDGFLTCVPGNASSGPQOLPIN 498  
Db 626 OKLAENHVTLPDPSLLVYRITDAASSWFDLGIN 658

## RESULT 10

C48562  
coat protein - San Miguel sea lion virus (serotype 4)  
N:Alternate names: capsid protein  
C:Species: San Miguel sea lion virus  
C:Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
C:Accession: C48562  
R:Neill, J.D.  
Virus Res. 24, 211-222, 1992  
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus  
A:Reference number: A48562; MUID:92410750; PMID:1529644  
A:Accession: C48562  
A:Molecule type: genomic RNA  
A:Residues: 1-703 <NEI>  
A:Cross-references: GB:M87482; NID:G334886; PIDN:AAA16220.1; PID:G334888  
A>Note: sequence extracted from NCBI backbone (NCBI:113567)  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted  
Query Match 8.9%; Score 251; DB 1; Length 703;  
Best Local Similarity 25.1%; Pred. No. 6.4e-10;  
Matches 116; Conservative 67; Mismatches 188; Indels 92; Gaps 23;  
QY 13 DGASGA-----GOLVPEVNASDDPLAMPVAGSSAVATAGQVNPIDPWI-----I 57  
Db 154 DPGSABEIVTEEQGTVVOQPPAPAPTALATLATASTGKSVEQE-----WMTFFSVHTSI 207  
QY 58 NNFOAPGGEFTISPNTPGGVLPDLGLPHLNPFLHLSQMYNGVGNMVRIMLAGNA 117  
Db 208 NMSTVESQCK-----ILYSQALNPSPINPYLDHIAKLYSTWSGGIDVFTVSGSG 256  
QY 118 FTAGKIIIVSCIPPGFGS-HNLTIAQATLPHVIADVRLDPIEVLVDVRLVFNHNDNRN 176  
Db 257 VFGGKLAALLVPPGVEPIESVMLQ---YPHVLFDARQTEVFIFIPDIRKTLFHSMDRT 313  
QY 177 QOTMLVCMLYTLPRLTGGTGDGDSFVVGVRVMTCPSPDFNLFPLVPP--TVEQKTRPFTL 234  
Db 314 -DTTKLVINPYE-----NGVENKTTCSITVETRPASDFTFALLKPPGSLKHGSLPSDL- 366  
QY 235 NLPLSSL--SNSRAPLPISGMGISPDNVQSVQFQNGRCTLDGRLVG-TTPVSLSHVAKIR 291  
Db 367 -IPRNSAHMGNRRWSTISGFSVQPRVFQS-----NRHFDFFDSTTTGWSPTYIPIEIKI 421  
QY 292 GT--SNGTVINLTDELDTGTFHPFEGPAPIGPPDLGGCDWHINMTQFQHSQTYQYDVTTP 349  
Db 422 GKVGNKNKWFHVIDTD-----KALVGPIDGWPDTTIPD-ETKATNGNFSYGESYRAGST- 475

QY 350 DTFVPHLSIQANGICGNYI-GVLSWVSPSPH-----PSGSQVDLWKIPNYGSSI 399  
 Db 476 -TIKENENSTH----KGTICIGTSTVEIPENDSQIKTEAEKKSQTMVYVVTADFKDTI 530  
 QY 400 TEATH-LAPSVYPPGGEVLVFF-----MSKIPGPGAYSL 433  
 Db 531 VKPQHKISQ-----KLWYFDGPEKDLTMSATLSPLGYTL 566

RESULT 11  
 JQ2354  
 coat protein - feline calicivirus (strain NADC)  
 C:Species: feline calicivirus  
 C>Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
 C:Accession: JQ2354  
 R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.  
 J. Gen. Virol. 74, 2519-2524, 1993  
 A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable  
 A:Reference number: JQ2354; MUID:94065683; PMID:7504075  
 A:Accession: JQ2354  
 A:Molecule type: mRNA  
 A:Residues: 1-668 <SEA>  
 A:Cross-references: GB:109718; NID:g305104; PIDN:AAA16485.1; PID:g305105  
 C:Superfamily: feline calicivirus coat protein  
 C:Keywords: capsid protein; coat protein

Query Match 8.6%; Score 242.5; DB 2; Length 668;  
 Best Local Similarity 32.2%; Pred. No. 2.4e-09;  
 Matches 64; Conservative 27; Mismatches 83; Indels 25; Gaps 6;  
 QY 34 DPVAGSSTAVATAGQVNPDPWLNINNVQAPQGEFTI-----SPNTPGGVLPDLSLGP 87  
 Db 145 EPSAQMSAADMATGSKVDSEW-----EAFFSHTSVNWTSETQKILFKQSLGP 195  
 QY 88 HLNPFLLHLSQMYNGVGNMRVIRIMLAGNAFTAGKIIVSCIPPGFSGHNLITIAQATLFP 147  
 Db 196 LLNPYLEHLKSLYVAMSGSVFVSISGSGVFGKLAIVVPPGVDPVQST--SMLQYPH 253  
 QY 148 VIADVRTLDPIEVLDRVNLPHNDRNQOTMRLVCMLY-----TPLRTGGGTGDSFVVA 203  
 Db 254 VLFDAQVDPVIFSPIDRLSTLYHLMF-DTDTTSLVIMVYNDLINPYANDSNSSGCIVT- 311

RESULT 12  
 VCWVFC  
 coat protein - feline calicivirus (strain Japanese F4)  
 N:Alternate names: capsid protein  
 C:Species: feline calicivirus  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 16-Jun-2000  
 C:Accession: B40481  
 R:Tohya, Y.; Taniguchi, Y.; Takahashi, E.; Utagawa, E.; Takeda, N.; Miyamura, K.; Yamaza  
 Virology 183, 810-814, 1991  
 A:Title: Sequence analysis of the 3'-end of feline calicivirus genome.  
 A:Reference number: A40481; MUID:91306470; PMID:1853578  
 A:Accession: B40481  
 A:Molecule type: genomic RNA  
 A:Residues: 1-668 <TOH>  
 A:Cross-references: GB:D90357; NID:g221264; PIDN:BAAL4371.1; PID:g221266  
 C:Superfamily: feline calicivirus coat protein  
 C:Keywords: capsid protein; coat protein; glycoprotein  
 C:Keywords: capsid protein; coat protein; glycoprotein (Asn) (covalent) #status predicted  
 F:177,301,304,399,459,615/Binding site: carbohydrate (Asn)

Query Match 8.6%; Score 241.5; DB 1; Length 668;  
 Best Local Similarity 31.7%; Pred. No. 2.8e-09;  
 Matches 63; Conservative 28; Mismatches 83; Indels 25; Gaps 6;  
 QY 34 DPVAGSSTAVATAGQVNPDPWLNINNVQAPQGEFTI-----SPNTPGGVLPDLSLGP 87  
 Db 145 EPSAQMSAADMATGSKVDSEW-----EAFFSHTSVNWTSETQKILFKQSLGP 195

QY 88 HLNPFLLHLSQMYNGVGNMRVIRIMLAGNAFTAGKIIVSCIPPGFSGHNLITIAQATLFP 147  
 Db 196 LLNPYLEHLKSLYVAMSGSVFVSISGSGVFGKLAIVVPPGVDPVQST--SMLQYPH 253  
 QY 148 VIADVRTLDPIEVLDRVNLPHNDRNQOTMRLVCMLY-----TPLRTGGGTGDSFVVA 203  
 Db 254 VLFDAQVDPVIFSPIDRLSTLYHLMF-DTDTTSLVIMVYNDLINPYANDSNSSGCIVT- 311  
 QY 204 GRVMTCPSPDPNLFVLP 222  
 Db 312 --VETKPGDPKFKHLLKPP 328

RESULT 13  
 VCWVFF  
 coat protein - feline calicivirus (strain CFI/68 FIV)  
 N:Alternate names: capsid protein  
 C:Species: feline calicivirus  
 C>Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 20-Aug-1999  
 C:Accession: A40507; B40507; T09246  
 R:Neill, J.D.; Reardon, I.M.; Heinrichson, R.L.  
 J. Virol. 65, 5440-5447, 1991  
 A:Title: Nucleotide sequence and expression of the capsid protein gene of feline caliciv  
 A:Reference number: A40507; MUID:91374597; PMID:1716692  
 A:Accession: A40507  
 A:Molecule type: genomic RNA  
 A:Residues: 1-668 <NEI>  
 A:Cross-references: GB:W32819; NID:g323874; PIDN:AAA42925.1; PID:g323875  
 A:Accession: B40507  
 A:Molecule type: protein  
 A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>  
 R:Neill, J.D.  
 submitted to the EMBL Data Library, April 1998  
 A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.  
 A:Reference number: Z16626  
 A:Accession: T09246  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: genomic RNA  
 A:Residues: 1-668 <NE3>  
 A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256  
 A:Experimental source: strain CFI/68 FIV  
 C:Superfamily: feline calicivirus coat protein  
 C:Keywords: capsid protein; coat protein; glycoprotein  
 F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted  
 F:177,301,304,439,459,615/Binding site: carbohydrate (Asn)

Query Match 8.5%; Score 240.5; DB 1; Length 668;  
 Best Local Similarity 29.9%; Pred. No. 3.3e-09;  
 Matches 64; Conservative 29; Mismatches 86; Indels 35; Gaps 7;  
 QY 19 GOLVPEVNASDPLAMPVAGSSTAVATAGQVNPDPWLNINNVQAPQGEFTI-----SP 72  
 Db 140 GGVIAPNAQMSAADMATGSKVD-----SEW-----EAFFSHTSVNWT 180  
 QY 73 NNTPGGVLPDLSGLPHLPFLHLSQMYNGVGNMRVIRIMLAGNAFTAGKIIVSCIPPGF 132  
 Db 181 SETQKILFKQSLGPLLPILYTHLAKLYVAMSGSVDFVSISGSGVFGKLAIVVPPGI 240  
 QY 133 GSHNLITIAQATLFPVIAVDVRLDPIEVLDRVNLPHNDRNQOTMRLVCMLY-----T 188  
 Db 241 DPVQST--SMLQYPHVLFDAQVDPVIFSPIDRLSTLYHLMF-DTDTTSLVIMVYNDLIN 297

RESULT 14  
 VCWVFF9  
 coat protein - feline calicivirus (strain F9)  
 N:Alternate names: capsid protein  
 C:Species: feline calicivirus  
 C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 28-Jul-2000  
 QY 189 PLRTGGGTGDSFVAGRVMTCPSPDPNLFVLP 222  
 Db 298 PYANDSNSSGCIVT---VETKPGDPKFKHLLKPP 328



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.  
OM protein - protein search, using sw model  
Run on: June 1, 2004, 13:34:56 ; Search time 6.40268 Seconds  
(without alignments)  
4310.252 Million cell updates/sec  
Title: US-09-926-799-2  
Perfect score: 2819  
Sequence: 1 MMASKDATSSVDGASGAGQ.....YQLKPVGTASSARGRLGLRR 530  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5  
Searched: 141681 seqs, 52070155 residues  
Total number of hits satisfying chosen parameters: 141681  
Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries  
Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2033.5	72.1	546	1 COAT_SOUV3	Q04542 southampton
2	1150.5	40.8	539	1 COAT_LORDV	P54635 lordsdale v
3	320.5	11.4	2344	1 POLN_RHDV	P27410 rabbit hemo
4	261.5	9.3	702	1 COAT_SMSV1	P36284 san miguel
5	252	8.9	2208	1 POLN_MANCV	Q69014 manchester
6	251	8.9	703	1 COAT_MNSV4	P36285 san miguel
7	241.5	8.6	668	1 COAT_FCVF4	P27405 feline cali
8	240.5	8.5	668	1 COAT_FCVF6	P27404 feline cali
9	237	8.4	671	1 COAT_FCVF9	P27406 feline cali
10	149	5.3	2205	1 POLG_POL2W	P23069 p genome po
11	147	5.2	2207	1 POLG_POL2L	P06210 p genome po
12	142	5.0	2194	1 POLG_HE701	P32537 h genome po
13	140.5	5.0	2206	1 POLG_POL32	P06209 poliovirus
14	138	4.9	2153	1 POLG_HRV16	Q82122 h genome po
15	136.5	4.8	2185	1 POLG_SVDVU	P13900 s genome po
16	134.5	4.8	2164	1 POLG_HRV89	P07210 h genome po
17	132	4.7	2193	1 POLG_HE71M	Q66478 h genome po
18	131	4.6	2193	1 POLG_HE71B	Q66478 h genome po
19	130.5	4.6	2432	1 Y43P_IRV6	P18305 chilo iride
20	128.5	4.6	2185	1 POLG_SVDVH	P18604 s genome po
21	128	4.5	855	1 POLG_HRV3	Q82081 human rhino
22	127.5	4.5	2175	1 POLG_BOVEV	P12915 b genome po
23	126.5	4.5	2196	1 POLG_EC05N	Q9V1J1 e genome po
24	126	4.5	2185	1 POLG_CXB5P	Q30353 c genome po
25	126	4.5	2206	1 POLG_CXA21	P22055 c genome po
26	126	4.5	2206	1 POLG_POL1M	P03299 p genome po
27	124.5	4.4	1324	1 SALI_HUMAN	Q9nscc2 homo sapien
28	123.5	4.4	862	1 POLG_EC16H	Q66790 echovirus 1
29	123	4.4	1200	1 HYAL_STRPU	P07636 strongyloce
30	122	4.3	2206	1 POLG_POL3L	P03302 poliovirus
31	122	4.3	2208	1 POLH_POL1M	P03300 p genome po
32	120.5	4.3	1183	1 DRPL_RAT	P54258 rattus norv
33	120	4.3	477	1 VL2_HPV42	P27235 human papil

34 119.5 4.2 2214 1 POLG\_CXA24 P36290 c genome po  
35 117 4.2 2179 1 POLG\_HRV14 P03303 h genome po  
36 116.5 4.1 832 1 POLG\_HRV1A P23008 human rhino  
37 116 4.1 2209 1 POLG\_POL1S P03301 p genome po  
38 114.5 4.1 2184 1 POLG\_EC01P O91734 e genome po  
39 113.5 4.0 592 1 YTRP\_PSEPU P40604 pseudomonas  
40 113.5 4.0 2688 1 ZEP1\_MOUSE Q03172 mus musculu  
41 113 4.0 1853 1 CIPA\_CLOTM Q06851 clostridium  
42 112 4.0 2194 1 POLG\_EC30B Q9wn78 e genome po  
43 111.5 4.0 2150 1 POLG\_HRV2 P04936 h genome po  
44 111.5 4.0 2157 1 POLG\_HRV1B P12916 h genome po  
45 111 3.9 2191 1 POLG\_EC06C Q66474 e genome po

ALIGNMENTS

RESULT 1  
COAT\_SOUV3  
ID COAT\_SOUV3 STANDARD; PRT; 546 AA.  
AC Q04542;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE Coat protein (Capsid protein).  
OS Southampton virus (serotype 3).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=37129;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93142023; PubMed=8380940;  
RA [1] PubMed=93142023; PubMed=8380940;  
RT "Sequence and genome organization of a human small round-structured (Norwalk-like) virus."  
RL [1] Sequence 2591836-510 (1999)  
CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
EMBL; L07418; AAA92984.1; -;  
PIR; B37491; B37491.  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
KW Coat protein; Glycoprotein.  
FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;  
Query Match 72.1%; Score 2033.5; DB 1; Length 546;  
Best Local Similarity 69.6%; Pred. No. 8.9e-139;  
Matches 380; Conservative 67; Mismatches 83; Indels 17; Gaps 6;  
QY 1 MMASKDATSSVDGASGAGQLVPEVNASDDPLAMPVAGSSTAVATAGQVNPIDPWVNF 60  
Db 1 MMASKDAQSADGASGAGQLVPEVNTADPLMPSPVAGPTTAVATAGQVNPIDPWVNF 60  
QY 61 VQAPQGETTISPNNTPGGVLFDLGLPHNLFLHLSQMYNGWGNMVRVRLMAGNAFTA 120  
Db 61 VQSPQGETTISPNNTPGDILFDLGLPHNLFLHLSQMYNGWGNMVRVRLMAGNAFTA 120  
QY 121 GKIIIVSCIPPGFGSHNLTAQATLPPHVIADRTLDPIEVLDRVNLPHNDRNQOTM 180  
Db 121 GKIIIVCCVPPGFTSSSLTQAQATLPPHVIADRTLEIEMPLEDRVNLVYHTND-NQPTM 179

		Matches 249; Conservative 82; Mismatches 182; Indels 39; Gaps 14;	
QY	181	RLVCMYTLPTLRTGGTGT--DSFVAVAGVMTCPSPDFNLFVLPPTVEQKTPFTLPL 238	
DB	180	RLVCMYTLPTLRTGGSGNSDSFVAVAGVLTAPSSDFSLFVLPPTIEKTRAFVPEIPL 239	
QY	239	SSLSNSRAPLPTISGMGIGPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRGTSN--G 296	
DB	240	QTLNSNRPSLSIQMLSPDASQVVFQNGRCLIDQLLGTTPATSGQLFVRGKINQGA 299	
QY	297	TVINILTELDGTFPHFPEGPAPGFDLGGCDWHNNMTQFHSS-----QTQVDVDTTPTD 351	
DB	300	RTNLNTEVDGKPFMAFDSAPVGFDFGKCDWHNNMTQFHSS-----QTQVDVDTTPTD 359	
QY	352	FVPHLGSIQANGI--GSGNYIGVLSWSPSPHSPSGQVDLWKIPNYGSSITEATHLAPS 408	
DB	360	FVPHLGSIQDFVENHPTGDYGTIEWISQPTPTGTDINLWEIPDYGSSLSQAANLAPP 419	
QY	409	VYPPGGEVLVFMKIPGFGAYS-----LPCLLPQEIYISHLASEQAPTVEAALHYVDP 464	
DB	420	VFPFGGEALVFPVSAFPPGNRSAPNDVFCLLPQEIYIFHVSQAPTMGDIALHYVDP 479	
QY	465	DTGRTIGEPKAYPDGFLTCVPNGASSGPQOLPINGVVFVSVSVRFYQLKPYGTASSARG 524	
DB	480	DTNRNLGEEKLYEGGYLTCPVNGVAGAPQOLPLNGVFLFVSVSVRFYQLKPYGTASTARG 539	
QY	525	RLGLRR 530	
DB	540	RLGVR 545	
RESULT 2			
ID	COAT LORDV	STANDARD; PRT; 539 AA.	
AC	P54635;		
DT	01-OCT-1996 (Rel. 34, Created)		
DT	01-OCT-1996 (Rel. 34, Last sequence update)		
DT	01-NOV-1997 (Rel. 35, Last annotation update)		
DE	Coat protein (Capsid protein).		
OS	Lordsdale virus (Human enteric calicivirus).		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OX	Norovirus.		
NCBI_TaxID=82658;			
RN	[1]		
RP	SEQUENCE FROM N.A.		
RP	MEDLINE=96005060; PubMed=7561776;		
RA	Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;		
RT	"Human enteric Caliciviridae: the complete genome sequence and		
RT	expression of virus-like particles from a genetic group II small		
RT	round structured virus.";		
RL	J. Gen. Virol. 76:2349-2355 (1995).		
CC	- - SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.		
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration		
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -		
CC	the European Bioinformatics Institute. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>		
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).		
CC	EMBL; X86557; CAA60255.1;		
DR	InterPro; IPR004005; Calici.coat.		
DR	InterPro; IPR008975; Viral_cap_coat.		
DR	Pfam; PF00915; Calici_coat; 1.		
KW	Coat protein; Glycoprotein.		
FT	CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).		
FT	CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).		
SQ	SEQUENCE 539 AA; 58775 MW; 5E5C63E7F2C5FD21 CRC64;		
Query Match 40.8%; Score 1150.5; DB 1; Length 539;			
Best Local Similarity 45.1%; Pred. No. 2.6e-75;			



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```

CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announcement/
CC or send an email to license@isb-sib.ch).
```

-----

EMBL; M67473; AAA47285.1; -

DR PIR; A41039; BRWWEH

DR PDB; 1KHV; 16-JAN-02.

DR PDB; 1KHV; 16-JAN-02.

DR MEROPS; C24.001; -

DR InterPro; IPR003593; AAA ATPase.

DR InterPro; IPR004005; Calici\_coat.

DR InterPro; IPR004004; Calici\_pol\_hel.

DR InterPro; IPR000317; Peptidase\_C24.

DR InterPro; IPR000605; RNA\_helicase.

DR InterPro; IPR007095; RNA\_pol\_DS\_PS.

DR InterPro; IPR001205; RNA\_pol\_P3D.

DR InterPro; IPR007094; RNA\_pol\_P5vir.

DR InterPro; IPR008975; Viral\_cap\_coat.

DR Pfam; PF00915; Calici\_coat; 1.

DR Pfam; PF03510; Peptidase\_C24; 1.

DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.

DR Pfam; PF00910; RNA\_helicase; 1.

DR PRINTS; PR00916; 2CENDOPTASE.

DR PRINTS; PR00918; CALICIVIRUS.

DR SMART; SM00382; AAA; 1.

DR Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;

KW Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.

FT CHAIN 7508 7707

FT CHAIN 1109 1251

FT CHAIN 71497 71625

FT CHAIN ? 2344

FT NP\_BIND 522 529

FT ACT\_SITE 1212 1212

FT ACT\_SITE 1227 1227

FT ACT\_SITE 2344 2344

FT SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212.CRC64;

Query Match 11.4%; Score 320.5; DB 1; Length 2344;

Best Local Similarity 24.2%; Pred. No. 8.4e-15;

Matches 136; Conservative 74; Mismatches 203; Indels 149; Gaps 24;

-----

4 ASKDATSSVDGASGAGQLVPEVNASDPLAMD- - - - -VAGSSTAVATAG- - - - - 47

1779 ACTATTASVPGTTD- - - - -MDPGVATTSVITAENSASIAATAGTGGPPQ 1825

-----

48 QVNPIDPWIINNFVQAPQGEFTISPNTPGGVLFDSLGLPHLNPFLHLSQMYNGVGNM 107

1826 QVQDQETWETNY- - - - -NDVFTWSVADAPGSLTYVQHSQPNPPTAVLSQMYAGWGM 1882

-----

108 RVRIMLAGNAFTAGKIIIVSCIPPFGF-SHNLTIQAATLPPHVIADVRLDPIEVLPLEVR 166

1883 QPRFIVAGSGVGGRLVRAVIPPGEIIGPLEVRQ- - - - -FPVVVIDARSLEPVTITMPDLR 1939

-----

167 NVLFHNDRNQMTMLVCMYLTPLRTGGTGDGDFVAGRVMTCPSPDFNLFLEVP- - - - -T 223

1940 PMMYHTGDPGLVPTLVLSVNNLINPFG-GSTSAIQVTVTRPSEDEFVIMRAPSKT 1998

-----

224 VEOKTRPFTLPNPLSSLSN-SRAPLPSGMGIPDNVQSVQFGNRCRTLDCRLVGTTPV 282

1999 VDSISGAGLLTTPVTGVDNRNNGQIVGLQVPGGFSTC- - - - -NRHNLNGSTVGNSSP 2055

-----

283 SLSHVAKIRGTSNGTINLT- - - - -DGT- - - - -FHPFEGPAPI 318

2056 REGDIDHRRGSASYSGNATNLVQFWYANAGSAIDNPISQVAPDGFDMSPFVFN- - - - - 2112

-----

319 GPFDLG- - - - -GCDWHINTQFGHSQTQYDVDTTDPDTPVPHLGSQANGISGNVIGVLS 374

2113 GIPAGAGWVGFAIWSN- - - - -SCAPNTTVQAYELGPAT- - - - -GAPG 2151

-----

375 WYSPSPHSGSGVDLWKIPNYGSSITEATHLAPSVY- - - - -PPGFGEVLVFMASK- 424

2152 NLQPTINTSGAT- - - - -VAKSIYAVVTGTAQNAG- - - - -LRFVMSG 2189

-----

QY 425 -IPGCGAYSLCL-LPQBYISHLASEQAPTVCERALHLYVD- - - - -PDTGRTLGEFKAYP 477

DB 2190 IISTENASAITVTPQDRIVTPGTPAAAPVGKNTPIMFASVVRRTGVDVNATAGS- - - - - 2244

QY 478 DGFLTCVPNGA--SSGPQOLPI 497

DB 2245 - - - - -ANGTQYGTGSQPLPV 2259

-----

RESULT 4

COAT\_SMSV1

ID\_COAT\_SMSV1 STANDARD; PRT; 702 AA.

AC P36284;

DT 01-JUN-1994 (Rel. 29, Created)

DT 01-JUN-1994 (Rel. 29, Last sequence update)

DT 15-DEC-1998 (Rel. 37, Last annotation update)

DE Coat protein (Capsid protein).

OS San Miguel sea lion virus (serotype 1) (SMSV 1).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Vesivirus.

OX NCBI\_TaxID=36406;

RN SEQUENCE FROM N.A.

RP MEDLINE=92410750; PubMed=1529644;

RA Neill J.D.;

RT "Nucleotide sequence of the capsid protein gene of two serotypes of

RT San Miguel sea lion virus: identification of conserved and non-

RT conserved amino acid sequences among calicivirus capsid proteins.";

RL Virus Res. 24:211-223(1992).

CC -I- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

EMBL; M87481; AAA16217.1; -

DR PIR; A48562; A48562.

DR InterPro; IPR004005; Calici\_coat.

DR InterPro; IPR008975; Viral\_cap\_coat.

DR Pfam; PF00915; Calici\_coat; 1.

FW Coat protein; Glycoprotein.

KT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT SEQUENCE 702 AA; 77850 MW; E6E5A58523DEE3D7.CRC64;

Query Match 9.3%; Score 261.5; DB 1; Length 702;

Best Local Similarity 23.9%; Pred. No. 3e-11;

Matches 137; Conservative 59; Mismatches 222; Indels 155; Gaps 26;

-----

QY 13 DGASGAGQLVPE- - - - -VNASDPLAMDPLVAGSSTAVATAGQVNPIDPWIINNFVQAPQGEFT 69

DB 154 DPGGADIVTEEGTIVVQQFV- - - - -PAQSALTTLAAASTGKTVD- - - - -CEWT 198

-----

70 I- - - - -SPNTPGGVLFDSLGLPHLNPFLHLSQMYNGVGNVRIMLAGNAFT 119

199 TFFSYHTAVNMTTEAQKILFSRALSPELNPYLRHISLSYTWSSGGIDVRFVTSGSGVF 258

-----

120 AGKIIIVSCIPPFGS-HNLTIQAATLPPHVIADVRLDPIEVLPLEVRNLFHNDRNQ 178

259 GKKLAALIVPGIEPVESPTMLQ- - - - -YHVLFDARQETEVFTIPIDKTLVHSD-DTD 314

-----

179 TMRVLCMLYTP- - - - -RTGGTGDGDFVAGRVMTCPSPDFNLFLEVPVTEQKTRPFTL 233

315 TTRLVIMVYNELINPYEQSEPKSSCSITVETR- - - - -PSSDFTFSLKPKPGSLKHGSI 370

-----

234 PNLPLSS- - - - -LNSRAPLPISGMGISPDNVQSVQFGNRCRTLDCRLVGT- - - - -VSL 284

Db 371 DLIPRNSRHWGNRWSIDGFEVQVRFQS-----NRHFDSDSTTCGWSTPYIPIEVTIL 426  
 QY 285 SHV-----AKIRGTSNGTVINLTDLGTPHPPEGPAPIGFPPDL-----GGCDW 328  
 Db 427 EKLDRGGQVFKYVTDTEKSLVPL-----PDGWPDTTPTTAMTANGNYDY 471  
 QY 329 HI-----NMQF-----GHSSQT-----QYDVTTP 349  
 Db 472 TVAEVRIITNNGTHFGFYMGHLLTKVKGSDNLGETQTSRTLPASVGNKQDQNTINPTH 531  
 QY 350 DTFVPHLGSIQANGIGSGNYIGVLGVSPSPSPSGQVDLWKIPNGYSITEATHLA--P 407  
 Db 532 KITSNSLVVYDANNVSAAT-AKTTTWHSTMSHLGYLVVDESVP-----GSDSTKVVRATLP 587  
 QY 408 SVYPPGFGVLFVFFMSKIPGPGNYSILPCLLPQEYISHLASEQAPTVEAALLHYVDPDTG 467  
 Db 588 EAFNG-GNFPVFETTKI-----QGHFDRHAKTKFNSQVLM-----TS 625  
 QY 468 RTLGE--FRAYPDGFLTCVPNGASSGPQQLPIN 498  
 Db 626 QKLAENHYTLPPDSLIVYRITDAASSWFDLGIN 658

RESULT 5  
 ID POLN MANCV STANDARD; PRT; 2208 AA.  
 AC Q69014; 1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polyprotein [Contains: RNA-directed RNA polymerase  
 (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like  
 protein); Coat protein].  
 DE Manchester virus (human enteric calicivirus).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Sapovirus.  
 NCBI\_TaxID=82659;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=9539791; PubMed=7661689;  
 RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;  
 RT "Human enteric caliciviruses have a unique genome structure and are  
 distinct from the Norwalk-like viruses."  
 RL Arch. Virol. 140:1345-1356(1995).  
 CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).  
 CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR  
 CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY  
 CC SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA}(N).  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X86560; CAA60262.1; --  
 CC MEROPS; C24.UPW; --  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR004004; Calici\_pol\_hel.  
 DR InterPro; IPR00317; Peptidase\_C24.  
 DR InterPro; IPR00605; RNA\_helicase.  
 DR InterPro; IPR00703; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_PSVir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.

DR Pfam; PF03510; Peptidase\_C24; 1.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PR00916; 2CENDOPTASE.  
 DR PRINTS; PR00918; CALICIVIRUSNS.  
 KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;  
 KW Thiol protease; Helicase; ATP-binding; Coat protein.  
 FT CHAIN ? ? HELICASE (P2C-LIKE).  
 FT CHAIN ? ? PROTEASE (P3C).  
 FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.  
 FT CHAIN ? ? COAT PROTEIN.  
 FT NP\_BIND 408 415 ATP (POTENTIAL).  
 FT ACT\_SITE 1097 1097 PROTEASE (BY SIMILARITY).  
 FT ACT\_SITE 1112 1112 PROTEASE (BY SIMILARITY).  
 FT ACT\_SITE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;  
 SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;  
 Query Match 8.9%; Score 252; DB 1; Length 2208;  
 Best Local Similarity 24.3%; Pred. No. 6.5e-10; Indels 58; Gaps 11;  
 Matches 99; Conservative 57; Mismatches 193;  
 QY 17 GACQLVPEVNASDPLAMDVPVAGSSTAVATAGQVNPIDP-----WIINNFVQ 62  
 Db 1650 GNGSNPEPKOSNPMVVDPPGTTGTTSHVVANPEQPNGAAQRLLEAVATGAIOSNVPE 1709  
 QY 63 APOQEFTI-----SPNPTPGGVLFDLGLPHLNPFLHLSQMYNGVGNMVRIMLAGN 116  
 Db 1710 AIRNCPAVFRTFAWDRMPGTGTLGSLHNPINPYTSHLSGMWAGMGSGFEVRLSISG 1769  
 QY 117 AFTAGKIIVSCIRPGFSGHNLTAQATLPHVIAVRTLDPIEVPLEDVRLVFNHNDRN 176  
 Db 1770 GVFAGRITIASVIPPVVDPS--SIRDGVLPFAVDARITEPVSMIPDVRADVHRMDGA 1827  
 QY 177 QOTMRLVCMLYTLPLRTGGTGDGDFVWAGRVMTCPSPDFFNLFLVPP--TVEQKTRFTLP 234  
 Db 1828 EPTCSLGFVWYQPLLNPFTTAVSTCWVSVETKPGDGFDFCLLRPPCQOMENGVSPEGLL 1887  
 QY 235 NLPLSSLSNRAPLPISGMGISPDNVQ--SVQFQNGRCTLGRLVGTTPVLSLHVA--KIRG 292  
 Db 1888 PRRLGYRGNRVGLVGVGMILVAEHKQVNRHFNNSVTFGWSTAPVNPMAAEIVTNOAHS 1947  
 QY 293 TSNQGVINLTDLGTPHPPEGPAPIGFPD-----LGG-----CDWHINMT 333  
 Db 1948 TSHRWLSIGAKNGKPLP---GIPNHPDSCASTVVGAMDTSLGGRPSTGVCSPAISFQ 2004  
 QY 334 QGHSSQYQYVDVTTPTDTPFVPHLSIQANGISGNYIGVLVSWVSPPS 380  
 Db 2005 NNG-----DVYENDTFVNFATY-----DPLTSGTGVALTNSINPAS 2041  
 RESULT 6  
 ID COAT SMSV4 STANDARD; PRT; 703 AA.  
 AC P36285;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS San Miguel sea lion virus (serotype 4) (SMSV 4).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 NCBI\_TaxID=36407;  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92410750; PubMed=1529644;  
 RA Neill J.D.;  
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of  
 RT San Miguel sea lion virus: identification of conserved and non-  
 RT conserved amino acid sequences among calicivirus capsid proteins."  
 RL Virus Res. 24:211-222(1992).  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; M87482; AA16220.1; --  
 CC FIR; C48562; C48562.  
 CC InterPro; IPR004005; Calici\_coat.  
 CC InterPro; IPR008975; Viral\_Cap\_coat.  
 CC Pfam; PF00915; Calici\_coat; 1.  
 CC Coat protein; Glycoprotein.  
 CC FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC SEQUENCE 703 AA; 77721 MW; C5DAD8223B261073 CRC64;

Query Match 8.9%; Score 251; DB 1; Length 703;  
 Best Local Similarity 25.1%; Pred. No. 1.7e-10;  
 Matches 116; Conservative 67; Mismatches 188; Indels 92; Gaps 23;

QY 13 DGASGA-----GQLVPEVNASDPLAMPVAGSSTAVATAGQVNPIDPW-----I 57  
 DB 154 DPGSAEIVTEEGTVVQOQAPAPTALATLASTAGKSVQE-----WMTFFSVHTSI 207  
 QY 58 NNFVQAPQGEFTIPNNTPGVFLDLSGLPHLPFLHLSOMYNGVGNMVRIMLAGNA 117  
 DB 208 NMTSVESQK-----ILYSQALNPSINPYLDHIAKLYSTWSGGIDVRFVSGSG 256  
 QY 118 FTAGKIIVSICPPFGS-HNLTIAQATLPHVIADVRLDPIEVLPLEVRNLFHNDNRN 176  
 DB 257 VFGKLAALLVPPGVEPIESVMLQ---YPHVLFDAQTEPVIPTIPDIRKTLFHSMDET 313  
 QY 177 QOTMLVCMLYPLRTGGGTGSFVAGVMTCPSPDFNLFVLP--TVQKTRPFTLP 234  
 DB 314 -DTTKLVINPYE-----NGVENKTCITVETRPSADFTFALLKPGSLIKGSPSLL- 366  
 QY 235 NPLSSL--SNSRAPLPIGSMGISPDNVQSFQNGRCRTDGLRVG--TTPVSLSHVAKIR 291  
 DB 367 -IPRNSAHMGNWSTISGFSVQPRVFS-----NRHFDFTSTTGWSTPYVPIEIKQ 421  
 QY 292 GT--SNGTVINTELDTGTFHPPEGPAPIGFPLDGGCDWHNNTPQHSQTYQVDVTTTP 349  
 DB 422 GKVGSNKWFHVIDTD---KALVPGIPDGMPTTIPD--ETKATNGNFSYGESYRAGST- 475  
 QY 350 DTFVPHLGIQANGISGNYI-GVLSWVSPSH-----PSGQVDLWKIPNVGSSI 399  
 DB 476 -TIKPNENSTHF-----KGTIYICGLSTVEIPENDEQOIKTEAPKKSQTMVYVVTADPKOTI 530  
 QY 400 TEATH-LAPSVYPPGGEVLVFP-----MSKIPGPGAYSIL 433  
 DB 531 VRPQHKISQ-----KLIVTFDGPKEKDLTWSATLSPLGYTL 566

## RESULT 7

COAT\_FCVF4  
 ID COAT\_FCVF4 STANDARD; PRT; 668 AA.  
 AC F27405;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS Feline calicivirus (strain Japanese F4) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OX NCBI\_TaxID=11980;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91306470; PubMed=1853578;  
 RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,  
 RA Miyamura K., Yamazaki S., Mikami T.;

RT "Sequence analysis of the 3'-end of feline calicivirus genome.";   
 RL Virology 183:810-814 (1991).  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; D90357; BAAL4371.1; --  
 CC FIR; B40481; VCMWPC.  
 CC InterPro; IPR004005; Calici\_coat.  
 CC InterPro; IPR008975; Viral\_Cap\_coat.  
 CC Pfam; PF00915; Calici\_coat; 1.  
 CC Coat protein; Glycoprotein.  
 CC FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
 CC SEQUENCE 668 AA; 73589 MW; 85BDBC885804E503 CRC64;

Query Match 8.6%; Score 241.5; DB 1; Length 668;  
 Best Local Similarity 31.7%; Pred. No. 7.8e-10;  
 Matches 63; Conservative 28; Mismatches 83; Indels 25; Gaps 6;

QY 34 DPVAGSSTAVATAGQVNPIDPWIIINNVQAPQGEFTI-----SPNTPGGVFLDLSLGP 87  
 DB 145 EPSAQMAADWASGKSVDSW-----EAFESFHTSVNMWSTSETQGLFKQSLGP 195  
 QY 88 HLNPELLHLSOMYNGVGNMVRIMLAGNAPTAKIIVSICPPGFGSHNLTAQATLFPH 147  
 DB 196 LNPFLZHLKUKYVWMSGIEVRFSIGVGKLAIVFPGVDPVQST--SMLOIYH 253  
 QY 148 VIADVRLDPIEVLPLEVRNLFHNDNRNQOTMLVCMLY-----TPLRTGGGTGSFVVA 203  
 DB 254 VLFDARQVEPVFTIPDLRSTLYHVS-DTDTTSLVIMVYNDLINPYANDSSGCIIVT- 311  
 QY 204 GRVMTCPSPDFNLFVLP 222  
 DB 312 --VETKPGDFKFKLLKPP 328

## RESULT 8

COAT\_FCVG6  
 ID COAT\_FCVG6 STANDARD; PRT; 668 AA.  
 AC P27404;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OX NCBI\_TaxID=11979;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=91374597; PubMed=1716692;  
 RA Neill J.D., Reardon I.M., Heinrichson R.L.;  
 RT "Nucleotide sequence and expression of the capsid protein gene of  
 RT feline calicivirus";  
 RL J. Virol. 65:5440-5447 (1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Neill J.D.;  
 RA Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- PTM: The N-terminus is blocked.  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.  
 CC -----



CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-aib.ch/announce/>  
 CC or send an email to [license@isb-aib.ch](mailto:license@isb-aib.ch)).  
 CC -----  
 CC EMBL; D00625; BAA00516.1; AUT\_SEQ.  
 DR PIR; A34032; GNNY2W.  
 DR HSP; P03299; IPOV.  
 DR MEROPS; C03.001; -.  
 DR MEROPS; C03.020; -.  
 DR InterPro; IPR003593; AAA\_AtPase.  
 DR InterPro; IPR004004; Calici\_pol\_hel.  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR001139; Pept\_3C\_picorn.  
 DR InterPro; IPR000081; Peptidase\_C3.  
 DR InterPro; IPR003138; Pico\_P1A.  
 DR InterPro; IPR002527; Pico\_P2B.  
 DR InterPro; IPR001676; Rhv.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_P5vir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00548; Cys-protease\_3C; 1.  
 DR Pfam; PF02226; Pico\_P1A; 1.  
 DR Pfam; PF00947; Pico\_P2A; 1.  
 DR Pfam; PF01552; Pico\_P2B; 1.  
 DR Pfam; PF00073; rhv\_3.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PR00918; CALICVIRUSNS.  
 DR ProDom; PD001125; Cys\_protease\_3C; 1.  
 DR ProDom; PD001306; Pico\_P2A; 1.  
 DR ProDom; PD001274; Pico\_P2B; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PolyProtein; Coat protein; Core protein; Transferase;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 KW Lipoprotein.  
 FT CHAIN 2 69 COAT PROTEIN VP4.  
 FT CHAIN 70 340 COAT PROTEIN VP2.  
 FT CHAIN 341 578 COAT PROTEIN VP3.  
 FT CHAIN 579 879 COAT PROTEIN VP1.  
 FT CHAIN 880 1028 PROTEASE 2A.  
 FT CHAIN 1029 1125 CORE PROTEIN 2B.  
 FT CHAIN 1126 1454 CORE PROTEIN 2C.  
 FT CHAIN 1455 1541 CORE PROTEIN 3A.  
 FT CHAIN 1542 1563 GENOME-LINKED PROTEIN VP6.  
 FT CHAIN 1564 1746 PICORNAIN 3C.  
 FT CHAIN 1747 2205 RNA-DIRECTED RNA POLYMERASE 3D.  
 FT LIPID 2 2 N-myristoyl glycine (in host) (By  
 FT similarity).  
 FT ACT\_SITE 1710 1710 PROTEASE (POTENTIAL).  
 FT ACT\_SITE 1724 1724 PROTEASE (POTENTIAL).  
 SQ SEQUENCE 2205 AA; 245701 MW; 2A42AB039E0254AD CRC64;  
 Query Match 5.3%; Score 149; DB 1; Length 2205;  
 Best Local Similarity 19.8%; Pred. No. 0.017;  
 Matches 110; Conservative 79; Mismatches 227; Indels 140; Gaps 23;  
 QY 22 VPEVNASDPLAMPVAGSSTAVATAGVNPIDVNNFVQAPQGFSTISPN-----NTP 76  
 DB 365 IFEFDVTPPI---DIFGEVNNMELAEIDTMTPLNTLSQKTYMDMYRVELNDAHSDFP 421  
 QY 77 GGVFLDLSGLPHLNPFLH-----LSQMYNGWGNRVRIMLAGNAFTAGKIIVSCIPPG 131

Db 422 ----ILCLSLSPASDRLAHTMGEILNYTHWAGSLKFTFLFCGSMATKLLVSVAPPG 478  
 QY 132 FGSNHLTIAQATLFFPHVIADVRLDPIEVPLDVRNVLFFHNNDRNQOT-----MR 181  
 Db 479 AKAPB-SRKEAMGLTHVIWDIGLQSSCTMWVPWISNTTYRQTINDSFTGGVISMFYQTR 537  
 QY 182 LVCMYLTPLRTGGTGGDSFVAVAGVMTCPSPFNELF-----LVPPT 223  
 Db 538 VVPLSTPRKMD-----ILGFVSAC--NDFSVRLLRDTHISQEVNPOGLDLEGV 587  
 QY 224 VEQKTRPFTPLNPLSLSSNSRAPLPISGMGTSPONVQSVQFQNGRCTLDGRLVGTTPVS 283  
 Db 588 VEGVTRNALPTLPVNNLPDTRS-----SGPAHSKETPALTAVETG---ATNPLVPSDTVQ 640  
 QY 284 LSHVAKIRGTSNGTV-----INLTDLGCTPFHPFEGPAPIGFPDLDGGCDWHIN--- 331  
 Db 641 TRHVIQKTRSESTVESFFAGACVAIIED-----NDAPTRASKLFSVWKITYKD 692  
 QY 332 ----MTQFGHSSQTYDVTDPDTFVPHLGSIQANGISGNVIGLVSWVSPSHSGSOV 387  
 Db 693 TVQLRRKLEFFTYSRDFMEF---TEFVTSNYTDANNHALNQVYQIMYI-PFGAP----- 743  
 QY 388 DLWKIPNYGSSITEATHLAPSVY-----PP-----GFGSVLVFFNS 423  
 Db 744 ----IPGRKNDVTWQTSNPSVFYTYGAPPARISVPYVIGIANAYSHFYDGFARVPLAGQA 799  
 QY 424 KIPGPGAYSLPCLLPQEYIS---HLASEQAPTVGBAALLHYVDVDPDTGRT----- 469  
 Db 800 STEGDSLYCAASL--NDFGSLAVRVVNDHNFTKLSKIRVYMKPKHVRWCPRPRAVPY 857  
 QY 470 LGFEKAYDPDFLTCTVP 485  
 Db 858 YGPGVDYKDG-LTPLP 872  
 RESULT 11  
 POLG\_POL2L  
 ID POLG\_POL2L STANDARD; PRT; 2207 AA.  
 AC P06210;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2  
 DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Picornain 2A  
 DE (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein  
 DE P3A; Genome-linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28)  
 DE (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].  
 OS Poliovirus type 2 (strain Lansing).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 OK NCBI\_TaxID=12084;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=86115399; PubMed=3003384;  
 RA la Monica N., Meriam C., Racaniello V.R.;  
 RT "Mapping of sequences required for mouse neurovirulence of poliovirus  
 RT type 2 Lansing."  
 RL J. Virol. 57:515-525 (1986).  
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
 CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.  
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the  
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
 CC substituted for Gln, and Ser or Thr for Gly.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC [RNA] (N).  
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; M12197; AAA46912.1; --  
 DR RIR; A29507; GNNVSP.  
 DR HSP; P03299; IPOV.  
 DR MEROPS; C03.001; --  
 DR MEROPS; C03.020; --  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR004004; Calici\_pol\_hel.  
 DR InterPro; IPR009003; Cys\_ser\_trypsin.  
 DR InterPro; IPR000199; Pept\_3C\_picorn.  
 DR InterPro; IPR000081; Peptidase\_C3.  
 DR InterPro; IPR003138; Pico\_P2A.  
 DR InterPro; IPR002527; Pico\_P2B.  
 DR InterPro; IPR001676; Rhv.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_PSVir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00548; Cys-protease\_3C; 1.  
 DR Pfam; PF02226; Pico\_P2A; 1.  
 DR Pfam; PF00947; Pico\_P2A; 1.  
 DR Pfam; PF01552; Pico\_P2B; 1.  
 DR Pfam; PF00073; rhv; 3.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PD00918; CALICIVIRUSNS.  
 DR ProDom; PD001125; Cys\_protease\_3C; 1.  
 DR ProDom; PD001306; Pico\_P2A; 1.  
 DR ProDom; PD001274; Pico\_P2B; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR PolyProtein; Coat protein; Core protein; Transferase;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 KW Lipoprotein.  
 FT CHAIN 2 69 COAT PROTEIN VP4.  
 FT CHAIN 70 340 COAT PROTEIN VP2.  
 FT CHAIN 341 578 COAT PROTEIN VP3.  
 FT CHAIN 579 879 COAT PROTEIN VP1.  
 FT CHAIN 880 1028 PROTEASE 2A.  
 FT CHAIN 1029 1125 CORE PROTEIN 2B.  
 FT CHAIN 1126 1454 CORE PROTEIN 2C.  
 FT CHAIN 1455 1541 CORE PROTEIN 3A.  
 FT CHAIN 1542 1563 GENOME-LINKED PROTEIN VP6.  
 FT CHAIN 1564 1746 PICORNAIN 3C.  
 FT CHAIN 1747 2207 RNA-DIRECTED RNA POLYMERASE 3D.  
 FT LIPID 2 2 N-myristoyl glycine (in host).  
 FT ACT\_SITE 1710 1710 PROTEASE (POTENTIAL).  
 FT ACT\_SITE 1724 1724 PROTEASE (POTENTIAL).  
 FT SEQUENCE 2207 AA; 245829 MW; 2B1E2070B7D44F99 CRC64;

Query Match  
 Best Local Similarity 20.1%; Pred. No. 0.023; Indels 112; Gaps 20;  
 Matches 99; Conservative 71; Mismatches 210;

QY 22 VPEVNASDPLAMPVAGSGSTAVATAGQVNPIDPIMNFVQAPQGEFTISPN-----NTP 76  
 DB 365 IPEFDVTPPI-----DIPGEVRNMELAEIDTIPNLNTRQNTWDMYRVELNDAHSDTP 421  
 QY 77 GGVLFDLSGLPHNPLHL-----LSQMYNGVGNVRIMLAGNFTAGKIIVSCIPPG 131  
 DB 422 ---ILCLSLSPASDPLRAHTMLGEILNYTHWAGSLKFTFLFCGSMWATGKLVSYPGG 478  
 QY 132 FGSNLTIAQATLPHFVIADVRLDPIEVLVDNRNVLFNHNDNRQNT-----MR 181  
 DB 479 ASAPK-SRKEAMGLTHVINDIGLQSCSTWVWISNTYRQTINDSFTEGGVSMFYQTR 537  
 QY 182 LVCMLYTLRLTGGTGDSEFVAGRVMTCPSPDFNLF-----LVPPT 223

Db 538 VVVPISFPRKMD-----ILGFVSAC--NDFSVRLRLDTHISQEAAMPQGLDLIEGV 587  
 QY 224 VEQKTRPFTPLNPLUSSLSNSRAPLPISGMGISSPDNVQSVQFQNGRCTLGRLVGTTPVS 283  
 Db 588 VEGVTRNALTPPLTANNLPDTS-----SGPAHSKETPALTAVETG---ATNPLVPSDTVQ 640  
 QY 284 LSHVAKIRGTSGTV-----INLTDLGTTPHPFEGPAPIGFPDIPGDCDWHIN--- 331  
 Db 641 TRHVIQKRTRSESTVESFFARGACVAIIEDV-----NDAPTRKASLFSVMKITYKD 692  
 QY 332 ----MTQFGHSSQTYDVTTPDTFVPHLGSQAINGISGNYIGVLSWTSPPSHSGSQV 387  
 Db 693 TVQLRRKLEFFTYSRFDMEF---TFVVTISNYTDANNHALNQVYQIMYI-PPGAP----- 743  
 QY 388 DLWKIPNYGSSITEATHLAPSVYPGPGEVLPFEMSKIIPGPGAYSLPCLLPQEIYISHL-- 445  
 Db 744 ----IPGKWNIDYWTQTSNPS-----VFYTYGAP-PARISVYVYGIANAYSHFYD 788  
 QY 446 ASEQAPTVGEAA 457  
 Db 789 GFAKVPLAGQAS 800

## RESULT 12

POLG\_HE701 STANDARD; PRT; 2194 AA.  
 ID POLG\_HE701  
 AC P32537;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DE 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2  
 DE (PIB); Coat protein VP3 (PIC); Coat protein P2C; Core protein P2A; Genome-  
 DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
 DE linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
 DE (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
 OS Human enterovirus 70 (strain J670/71) (EV 70).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 OX NCBI\_TaxID=31915;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91037960; PubMed=2172447;  
 RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D.,  
 RA Minor P.D., Almond J.W.;  
 RT "The complete nucleotide sequence of enterovirus type 70:  
 RL relationships with other members of the picornaviridae.";  
 RL J. Gen. Virol. 71:2291-2299(1990).  
 CC -|- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
 CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.  
 CC -|- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the  
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
 CC substituted for Gln, and Ser or Thr for Gly.  
 CC -|- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC (RNA)(N).  
 CC -|- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -|- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -|- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; D00820; BAA18891.1; --  
 CC PIR; A36253; GNNYET.  
 CC HSPP; P03300; IPOV.  
 CC MEROPS; C03.00A; --

```

DR MEROPS; C03.UBP; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys_Ser_trypsin.
DR InterPro; IPR000199; Pept_3C_picorn.
DR InterPro; IPR000081; Peptidase_C3.
DR InterPro; IPR003138; Pico_P1A.
DR InterPro; IPR002527; Pico_P2B.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_Psvir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00548; Cys-protease_3C; 1.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00947; Pico_P2A; 1.
DR Pfam; PF01552; Pico_P2B; 1.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR PRINTS; PR00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR ProDom; PD001125; Cys_protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
DR PolyProtein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 319 COAT PROTEIN VP2.
FT CHAIN 320 561 COAT PROTEIN VP1.
FT CHAIN 562 871 COAT PROTEIN VP1.
FT CHAIN 872 1014 CORE PROTEIN P2A.
FT CHAIN 1015 1113 CORE PROTEIN P2B.
FT CHAIN 1114 1443 CORE PROTEIN P2C.
FT CHAIN 1444 1532 CORE PROTEIN P3A.
FT CHAIN 1533 1554 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1555 1737 PICORNAIN 3C.
FT CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 N-myristoyl glycine (in host) (By
similarity).
FT ACT_SITE 1701 1701 PROTEASE (POTENTIAL).
FT ACT_SITE 1715 1715 PROTEASE (POTENTIAL).
SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;

Query Match 5.0%; Score 142; DB 1; Length 2194;
Best Local Similarity 29.1%; Pred. NO. 0.052;
Matches 50; Conservative 22; Mismatches 84; Indels 16; Gaps 6;

QY 57 INNFOAQQE---FTISPNNTPGVLFDSL-----GPHLPFLHLQMYNGWVGNMR 108
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
374 INNVDASGERLRVQISQSDMDQLFNIPLDIQLEGPLRNTLLGNISRYVTHMSGSL 433
QY 109 VRIMLAGNFTAGKIIVSCIPGFGSHNITIAQTLFPHVIAD--VRTLDPIEVLVDVR 166
Db : : : : : : : : : : : : : : : : : : : : : : : : : : :
434 MTFMFCGSPWTKLIICYTPPG-GSSPTDRMQAMLATHVWDFGLQSSITIIIPWISGS 492
QY 167 NVLFHNDR---NQOTMRLVCMLYTLRTGGTGDSFVAGRVMTCPSPDN 215
Db : : : : : : : : : : : : : : : : : : : : : : : : : : :
493 HYRMFNTDAKAINANVGYVTCFMQTNLVAPVGAADQCIYV--MVAAKKDFN 542

RESULT 13
POLG_POL32
ID POLG_POL32 STANDARD; PRT; 2206 AA.
AC P06209;
DT 01-JAN-1988 (Rel. 06, Created)
DT 01-JAN-1988 (Rel. 06, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins
DE P2A TO P2C; P3A; Genome-linked protein VPG; Picornain 3C
DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D

```

```

DE (EC 2.7.7.48)].
OS Poliovirus type 3 (strain 23127).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12087;
RN [1]_TaxID=12087;
RP SEQUENCE FROM N.A.
RX MEDLINE=87010550; PubMed=3020156;
RA Hughes P.J., Evans D.M.A., Minor P.D., Schild G.C., Almond J.W.,
RA Stanway G.;
RT "The nucleotide sequence of a type 3 poliovirus isolated during a
RT recent outbreak of poliomyelitis in Finland.";
RL J. Gen. Virol. 67:2093-2102(1986).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC (RNA) (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; X04468; CAA28155.1; -.
CC PIR; A27245; GNNY27.
CC HSP; Q84790; LPVC.
CC MEROPS; C03.001; -.
CC MEROPS; C03.020; -.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000199; Pept_3C_picorn.
CC InterPro; IPR000081; Peptidase_C3.
CC InterPro; IPR003138; Pico_P1A.
CC InterPro; IPR002527; Pico_P2B.
CC InterPro; IPR001676; Rhv_P2B.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_Psvir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00548; Cys-protease_3C; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF00073; rhv; 3.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSNS.
CC ProDom; PD001125; Cys_protease_3C; 1.
CC ProDom; PD001306; Pico_P2A; 1.
CC ProDom; PD001274; Pico_P2B; 1.
CC SMART; SM00382; AAA; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 340 COAT PROTEIN VP2.
FT CHAIN 341 578 COAT PROTEIN VP3.
FT CHAIN 579 878 COAT PROTEIN VP1.
FT CHAIN 879 1027 CORE PROTEIN P2A.
FT CHAIN 1028 1124 CORE PROTEIN P2B.

```

FT CHAIN 1125 1453 CORE PROTEIN P2C.  
 FT CHAIN 1454 1540 CORE PROTEIN P3A.  
 FT CHAIN 1541 1562 GENOME-LINKED PROTEIN VPG.  
 FT CHAIN 1563 1745 PICORNAIN 3C.  
 FT CHAIN 1746 2206 RNA-DIRECTED RNA POLYMERASE P3D.  
 FT LIPID 2 N-myrystoyl glycine (in host) (By similarity).  
 FT ACT\_SITE 1709 1709 PROTEASE (POTENTIAL).  
 FT ACT\_SITE 1723 1723 PROTEASE (POTENTIAL).  
 SQ SEQUENCE 2206 AA; 245731 MW; F226AD85403C37BA CRC64;  
 Query Match 5.0%; Score 140.5; DB 1; Length 2206;  
 Best Local Similarity 20.4%; Pred. No. 0.068;  
 Matches 90; Conservative 67; Mismatches 186; Indels 99; Gaps 19;  
 QY 22 VEVNASDPLMDPVAGSTAVATAGQVNPIDPWIINNFPVQAPQGEFTI--SPNPTPGV 79  
 DB 365 IPEFDVTPPI--DIFGEVKNVMELEAIDTMIPLNLENTKRTMDMYRVLSDSANLSGP 421  
 QY 80 LPDLSLGPLNPLH---LSQYNGVGNVRVIRIMLAGNAFTAGKIIVSCIPPFGGS 134  
 DB 422 ILCLSLSPAADPLSLHTMLGEVLNYTHWAGSLKFTFLFCGSMATGKLIVAYAPFG-AQ 480  
 QY 135 HNTIAQAATLFPVIAVRLDIEVPLEDRVNLPHNDNRQQT-----MRLVC 184  
 DB 481 PPTSREKMLGTHVWDGLQSSCTMVVPMISNVTYRQTQDSFTGGVISMFYQTRIVV 540  
 QY 185 MLVTLRTGGTGDSFVAGRVMTCPSPDNFLFLVPPT-----VEQK 227  
 DB 541 PLSTP-----KAMDMLGFVSAC--NDFSVELLRDTHISQAAMPQGVDDLLITEVAQN 590  
 QY 228 TRPTPLNPLSLNSRAPLPISGMGISPDNVQSFQNGRCTLDRGLVGTTPVSLSHV 287  
 DB 591 ALALSLEPK-FQSNLPDPTKA---SGPAHSKEVETPLTAVETG---ATNPLVPSDTVQTRHV 642  
 QY 288 AKIRGTNGTV-----INLTDLGTPTFPFEGPAP-----IGFPDLGGCDW 328  
 DB 643 IQRSRSESTIESFFAGACVAIIEDVN-----EQATNVQKLFATWITYKDT--VQL 694  
 QY 329 HINMTQFCHSQTQYDVTTPDTPVPHLGSIQANGISGNYIGVLSVWSPSPHPSSQVD 388  
 DB 695 RKLEFFTYIS--RFDMEF---TFVTANFTNSNGHALNQVQIMYI-PPGAPT----- 742  
 QY 389 LMKIPNYGSSITEATHLAPSVY 410  
 DB 743 ---PKSMDDYTWQTSSNPSIF 760  
 RESULT 14  
 ID POLG\_HRV16 STANDARD; PRT: 2153 AA.  
 AC Q82122;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polypeptide [Contains: Coat protein VP4 (PIA); Coat protein VP2 (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
 OS Human rhinovirus 16 (HRV-16).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Rhinovirus.  
 OC NCBI\_TaxID=31708;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95250310; PubMed=7732663;  
 RA Lee W.M., Wang W., Rueckert R.R.;  
 RT "Complete sequence of the RNA genome of human rhinovirus 16, a clinically useful common cold virus belonging to the ICAM-1 receptor group.";  
 RL Virus Genes 9:177-181(1995).  
 RN [2]

RP X-RAY CRYSTALLOGRAPHY (3.5 ANGSTROMS) OF 2-853.  
 RX MEDLINE=94348864; PubMed=7915182;  
 RA Oliveira M.A., Zhao R., Lee W.M., Kremer M.J., Minor I.,  
 RA Rueckert R.R., Diana G.D., Pevear D.C., Dutko F.J., McKinlay M.A.,  
 RA Rossmann M.G.;  
 RT "The structure of human rhinovirus 16.";  
 RL Structure 1:51-68(1993).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.15 ANGSTROMS) OF 2-853, AND REVISIONS TO 547-548.  
 RP MEDLINE=97238938; PubMed=9083115;  
 RA Hadfield A.T., Lee W.M., Zhao R., Oliveira M.A., Minor I.,  
 RA Rueckert R.R., Rossmann M.G.;  
 RT "The refined structure of human rhinovirus 16 at 2.15-A resolution: implications for the viral life cycle.";  
 RL Structure 5:427-441(1997).  
 CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN O/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.  
 CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + [RNA] (N).  
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.  
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).  
 CC -----  
 CC EMBL; L24917; AAA69862.1; -.  
 DR PDB; 1AYN; 21-JAN-98.  
 DR PDB; 1AYM; 21-JAN-98.  
 DR MEROPS; C03.007; -.  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR004004; Calici\_pol\_hel.  
 DR InterPro; IPR009003; Cys\_ser\_tryptain.  
 DR InterPro; IPR001199; Pept\_3C\_picorn.  
 DR InterPro; IPR000081; Peptidase\_C3.  
 DR InterPro; IPR003138; Pico\_P1A.  
 DR InterPro; IPR002527; Pico\_P2B.  
 DR InterPro; IPR001876; RVV\_P2B.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_P3.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_PSVir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00548; Cys-protease-3C; 1.  
 DR Pfam; PF02226; Pico\_P1A; 1.  
 DR Pfam; PF00947; Pico\_P2A; 1.  
 DR Pfam; PF01552; Pico\_P2B; 1.  
 DR Pfam; PF00073; rhv; 3.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PR00918; CALICIVIRUSNS.  
 DR ProDom; PD001125; Cys\_protease\_3C; 1.  
 DR ProDom; PD001306; Pico\_P2A; 1.  
 DR ProDom; PD001274; Pico\_P2B; 1.  
 DR SMART; SM00382; AAA; 1.  
 KW Polypeptide; Coat protein; Core protein; Transferase;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 KW 3D-structure; Lipoprotein.  
 FT CHAIN 2 69 COAT PROTEIN VP4.  
 FT CHAIN 70 330 COAT PROTEIN VP2.  
 FT CHAIN 331 568 COAT PROTEIN VP3.  
 FT CHAIN 569 853 COAT PROTEIN VP1.



FT	CHAIN	854	995	CORE PROTEIN P2A.	FT	TURN	422	426
FT	CHAIN	996	1090	CORE PROTEIN P2B.	FT	TURN	428	433
FT	CHAIN	1091	1412	CORE PROTEIN P2C.	FT	TURN	434	435
FT	CHAIN	1413	1489	CORE PROTEIN P3A.	FT	STRAND	436	440
FT	CHAIN	1490	1510	GENOME-LINKED PROTEIN VPG.	FT	STRAND	443	449
FT	CHAIN	1511	1693	PICORNAIN 3C.	FT	TURN	453	454
FT	CHAIN	1694	2153	RNA-DIRECTED RNA POLYMERASE P3D.	FT	STRAND	456	456
FT	LIPID	2	2	N-myristoyl glycine (in host).	FT	STRAND	458	464
FT	ACT SITE	1657	1657	PROTEASE (POTENTIAL).	FT	STRAND	474	478
FT	ACT SITE	1671	1671	PROTEASE (POTENTIAL).	FT	TURN	479	479
FT	CONFLICT	547	548	KD -> NH (IN REF. 1).	FT	STRAND	481	486
FT	STRAND	27	30		FT	STRAND	492	497
FT	HELI	36	38		FT	STRAND	506	507
FT	TURN	80	81		FT	TURN	513	514
FT	STRAND	83	87		FT	STRAND	518	523
FT	TURN	88	89		FT	STRAND	528	528
FT	STRAND	90	94		FT	TURN	531	532
FT	STRAND	101	102		FT	STRAND	537	545
FT	HELI	103	105		FT	TURN	547	548
FT	TURN	113	115		FT	STRAND	550	554
FT	STRAND	123	123		FT	STRAND	567	567
FT	HELI	126	128		FT	HELI	570	579
FT	TURN	129	129		FT	TURN	580	580
FT	STRAND	133	134		FT	STRAND	584	584
FT	STRAND	138	140		FT	STRAND	588	588
FT	TURN	142	143		FT	TURN	600	601
FT	STRAND	147	151		FT	STRAND	602	603
FT	TURN	152	152		FT	HELI	605	607
FT	HELI	153	155		FT	HELI	615	618
FT	TURN	156	157		FT	STRAND	624	624
FT	HELI	159	167		FT	STRAND	629	629
FT	STRAND	168	180					
FT	TURN	185	186					
FT	STRAND	188	197					
FT	TURN	198	198					
FT	STRAND	204	204					
FT	TURN	207	208					
FT	HELI	213	216					
FT	HELI	219	221					
FT	STRAND	223	224					
FT	TURN	225	227					
FT	TURN	232	233					
FT	HELI	239	241					
FT	TURN	242	245					
FT	STRAND	246	246					
FT	HELI	248	253					
FT	STRAND	256	260					
FT	TURN	261	263					
FT	STRAND	266	271					
FT	STRAND	280	280					
FT	TURN	282	284					
FT	STRAND	285	285					
FT	STRAND	288	299					
FT	TURN	303	304					
FT	STRAND	307	323					
FT	TURN	338	341					
FT	STRAND	345	346					
FT	TURN	353	353					
FT	TURN	356	357					
FT	STRAND	369	370					
FT	STRAND	372	372					
FT	TURN	373	373					
FT	HELI	374	377					
FT	TURN	378	378					
FT	STRAND	381	382					
FT	TURN	385	386					
FT	STRAND	388	388					



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 46.9047 Seconds  
(without alignments)  
3289.030 Million cell updates/sec

Title: US-09-926-799-3

Perfect score: 2895

Sequence: 1 MMASKAPQADGASGAGQ.....QLKPVGTASTARSLGVRRRI 546

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_23Jan04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2895	100.0	546	4	AAB49702 Small rou
2	2231.5	77.1	544	4	AAB49703 Small rou
3	2034.5	70.3	530	2	AAR50972 Norwalk v
4	2034.5	70.3	530	7	Adc72176 Norwalk v
5	2029.5	70.1	530	4	AAB49701 Small rou
6	1993.5	68.9	530	2	AAR57091 Small rou
7	1977.5	68.3	545	4	AAB49700 Small rou
8	1175	40.6	535	4	AAB49707 Small rou
9	1169.5	40.4	540	4	AAB49706 Small rou
10	1155.5	39.9	542	4	AAB49708 Small rou
11	1124.5	38.8	548	4	AAB49705 Small rou
12	1115.5	38.5	548	5	AAU91272 Norwalk v
13	1105	38.2	539	4	AAB49704 Small rou
14	1090.5	37.7	550	4	AAB49709 Small rou
15	967	33.4	541	4	AAB49710 Small rou
16	321	11.1	579	2	AAW08143 RHDV caps
17	290	10.0	669	4	AAB67461 Amino aci
18	272	9.4	668	4	AAB67462 Amino aci
19	268.5	9.3	671	4	AAU50107 Feline ca
20	268	9.3	547	4	AAU50108 Feline ca
21	267	9.2	623	4	AAB47044 Feline Ca
22	267	9.2	623	4	AAB47043 Feline Ca
23	267	9.2	668	2	AAR10686 Feline ca
24	267	9.2	668	4	AAE04304 Feline ca
25	256	8.8	622	4	AAB47045 Feline Ca

26	144	5.0	2206	2	AAR22210	Aar22210 True type
27	134	4.6	40	5	AAU91273	AAU91273 Norwalk v
28	132.5	4.6	1194	6	ABU18239	ABU18239 Protein e
29	125.5	4.3	2209	1	AAP20037	Aap20037 Sequence
30	123	4.2	2735	6	ABR41356	ABr41356 Human DIT
31	123	4.2	2971	7	ADC35084	Adc35084 Human bre
32	123	4.2	2972	4	AAB50363	Aab50363 Human SRC
33	123	4.2	3118	4	AAB50362	Aab50362 Human SRC
34	122.5	4.2	6310	6	ABU39869	ABu39869 Protein e
35	121	4.2	40	5	AAU91274	AAu91274 Norwalk v
36	120	4.1	3443	2	AAW84559	AAw84559 Polyproti
37	119	4.1	2164	1	AAP81045	Aap81045 Sequence
38	118	4.1	2971	3	ABR41231	ABr41231 Human ORF
39	117	4.0	2164	1	AAP80131	Aap80131 Peptides
40	116.5	4.0	2150	1	AAP60045	Aap60045 Sequence
41	116.5	4.0	2150	1	AAP81048	Aap81048 Sequence
42	116.5	4.0	2150	2	AAO5127	Aar05127 Complete
43	116	4.0	2016	4	ABR63911	ABr63911 Drosophil
44	115.5	4.0	239	1	AAP81097	Aap81097 Sequence
45	115.5	4.0	1765	6	ABU20131	ABu20131 Protein e

## ALIGNMENTS

## RESULT 1

AAAB49702  
ID AAB49702 standard; protein; 546 AA.

XX AAB49702;

AC AAB49702;

DT 04-APR-2001 (first entry)

DB Small round structured virus protein SEQ ID 3.

KW Small round structured virus; SRSV; food poisoning.

OS Small round structured virus.

PN WO200079280-A1.

PD 28-DEC-2000.

PF 22-JUN-2000; 2000WO-JP004095.

PR 22-JUN-1999; 99JP-00175928.

PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

PA (DENK-) DENKA SEIKEN KK.

XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX WPI; 2001-080848/09.

XX N-PSDB; AAF29143.

XX Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

XX Claim 1; Page 45-47; 84pp; Japanese.

XX This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

XX Sequence 546 AA;

XX Query Match 100.0%; Score 2895; DB 4; Length 546;

XX Best Local Similarity 100.0%; Pred. No. 1.5e-257;

Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

Claim 1: Page 47-49: 84pp; Japanese.

This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49711, which are each SRSV strain specific. Polynucleotide sequences AAF20141, which are each SRSV strain specific, and polypeptide sequences AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the use of contaminated material in the epidemiology of outbreaks.

Sequence 544 AA;

DB 4:	DB 5:	DB 6:	DB 7:	DB 8:	DB 9:	DB 10:	DB 11:	DB 12:	DB 13:	DB 14:	DB 15:	DB 16:	DB 17:	DB 18:	DB 19:	DB 20:	DB 21:	DB 22:	DB 23:	DB 24:	DB 25:	DB 26:	DB 27:	DB 28:	DB 29:	DB 30:	DB 31:	DB 32:	DB 33:	DB 34:	DB 35:	DB 36:	DB 37:	DB 38:	DB 39:	DB 40:	DB 41:	DB 42:	DB 43:	DB 44:	DB 45:	DB 46:	DB 47:	DB 48:	DB 49:	DB 50:	DB 51:	DB 52:	DB 53:	DB 54:	DB 55:	DB 56:	DB 57:	DB 58:	DB 59:	DB 60:	DB 61:	DB 62:	DB 63:	DB 64:	DB 65:	DB 66:	DB 67:	DB 68:	DB 69:	DB 70:	DB 71:	DB 72:	DB 73:	DB 74:	DB 75:	DB 76:	DB 77:	DB 78:	DB 79:	DB 80:	DB 81:	DB 82:	DB 83:	DB 84:	DB 85:	DB 86:	DB 87:	DB 88:	DB 89:	DB 90:	DB 91:	DB 92:	DB 93:	DB 94:	DB 95:	DB 96:	DB 97:	DB 98:	DB 99:	DB 100:			
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

Query Match 77.1%; Score 196;  
Best Local Similarity 74.7%; Pred. NO. 2.2e-196;  
Matches 407; Conservative 59; Mismatches 78; Indels 1; Gaps 1;

QY  
1 MMMSKADPQSDAGSAGQLYPEVNTADPLPNEBPVAGPIYAVIAGQVNMDFEPIAKI 60

Db 1 MMASKDTPSADGATGAGQLVPEVNTADP IPIDFVAGSSAATRAHNOG+120

[illegible]

Db 61 VQAPQGEFTISFNNIPGVLFDFDQDQNNMTLSDNNL  
101 QKTHCCVCVDCGCTSSSTIAQATLPPHVIADRTLEPIEMPLDVRNVLVYHTNDNQPTWR 180

[illegible]

DB 121 GRVLLGGVLTGGRVLTAPSSDSEFLVPPTIEQKTRAFVTNIPLO 240

C= 181 IVCMLVTLPLRTGGGSGNSDFVAGRVLTAPSSDSEFLVPPTIEQKTRAFVTNIPLO 240

181 LLCLMYTPLRTGASGGTDSFVAGRVLTCPGPDFNLFVBPPTVEQKTRPFTVENIPLK 240

241 TLSNRFPESLIQGMILSPDASVVQVFQNGRCLIDGOLLGTTPTATSGQLFRVRGKINQGAR 300

db

301 TLNLTEVDGKPFMAFDSPAPVGFDFGKCDWHMRI SKTPNNNTSGDPMRSVSQITVQGF 360

Db 301 VLNLTELGSPFMAFAAPAPAGFPDLGSCDWHIEMSKIPNSSTQNPNPIVITNSVKPNSQQF 360

Qy 361 VPHLGSIQFDEVFNHPTGDYIGTIEWISQPSPPGTDINLWEIPDPYGSLSQAANLAPFV 4230

Db

361 VPHLSITLDENVSSGGDIYGTIQWTSPPSDGGANFNWKIPDYGSSLEAASQLPAFV 48

[illegible]

Db 420 YPPGFNEVIYFFMASTIPGPNQSGSPNLVFCLETFQEIILNFIISTEQNLNQSCT

481 TNRNLGEFKLYPGGYLFCVPGVGVAGPQQLFTNGTETKLSKQK 53  
QY

Db 480 TNRNLGEFKLIFGGGILLICVFNSSSICQVQZELSSSTF  
E43 YQMD E45

Qy	541 LGVRK 543
	540 LGVRD 544
z:	

DB 340 DEVARA 311

RESULT 3  
AAR50972

ID AAR50972 standard; protein; 530 AA.  
 XX

AC AAR50972;  
XX XX

DT	16-OCT-2003	(revised)
DT	25-MAR-2003	(revised)
	15-SEP-2004	(final entry)

DT	05-OCT-1994	(first entry)
XX		
XX	Newcastle virus strain APTia protein (encoded by ORF2).	

DE NOIWDIAA VIUUS DEBAM: CIIIT E  
XX

KW Norwalk virus; pathogen; acute gastroenteritis; food poisoning;  
KW seafood contamination; diagnostic assay; calicivirus; small round virus.  
OS Norwalk virus; (strain 8FIIa).  
XX WO9405700-A2.  
PN 17-MAR-1994.  
PD 07-SEP-1993; 93WO-US008447.  
XX 07-SEP-1992; 92US-00941365.  
XX (BAYU ) BAYLOR COLLEGE MEDICINE.  
PA Watson DO, Estes MK, Jiang X, Graham DY;  
XX WPI: 1994-101125/12.  
XX N-PSDB; AAQ56826.

DNA from Norwalk and related viruses - used for preparing prods. for use  
in diagnostic assays, detection and vaccines for Norwalk and related  
viruses.

Claim 14; Page 68-70; 156pp; English.

The Norwalk virus was isolated from stool samples from adult volunteers  
infected with safety tested Norwalk virus strain 8FIIa. The coding  
sequence is useful for the design of probes for use in diagnostic assays  
for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct  
PN field.) (Updated on 16-OCT-2003 to standardise OS field)

Sequence 530 AA;

Query Match 70.3%; Score 2034.5; DB 2; Length 530;  
Best Local Similarity 69.4%; Pred. No. 3.1e-178; Mismatches 17; Gaps 6;  
Matches 379; Conservative 69; Indels 17; Gaps 6;

QY 1 MMASKADPQSDAGSAGQGVPEVNTADPLPMEPVAGPTTAVATAGVNMIDPWVNNF 60  
DB 1 MMASKDATSSVDGASGAGQGVPEVNASDPLAMPVAGSSTAVATAGVNPDPWVNNF 60  
QY 61 VQSPGQETISPNNTPGDILFDLQGLPHNPLPFLSHLSQMYNGWGNMVRILLAGNAFSA 120  
DB 61 VQAPQGEFTISPNNTPGDVLFDLSGLPHNPLPFLSHLSQMYNGWGNMVRILLAGNAFTA 120  
QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPHVIAVTRTLEPIEMPLEDVNLYHTND-NQPTM 179  
DB 121 GKIIIVCCIPPPGGSHNLTIATLPPHVIADVRTLDPIEVLPLEDVNLYHTND-NQPTM 180  
QY 180 RLVCMLYTPLRGTGGSGNSDSFVWAGRVLTAPSSDFSLFLVPPTIEQKTRAFVNPIL 239  
DB 181 RLVCMLYTPLRGTGGTG--DSFVWAGRVMTCPSPDFNPLFLVPTVEQKTRFTLPNPL 238  
QY 240 QTLNSRFPFSLIQMILSPDASQVVOFQNGRCLIDGQLLGTTPATSGOLFVRGKINQGA 299  
DB 239 SLSLSNRAPLPISMGISPDVNVQSFQNGRCLTDLGRVLTGVTTPVSLSHVAKIRGTSN--G 296  
QY 300 RLNLNTEVDGKPFMAFDSPAPVGFDFGKCDWHMRISKTPNNTSGDPMRSVSUVTNVQ 359  
DB 297 TVINLTDLGTFHFHFEFGAPVGFDFGKCDWHMRISKTPNNTSGDPMRSVSUVTNVQ 351  
QY 360 FVPHLGSIQFDFVFNHPTGDIYGTIEWISQSPTEPGTDINLWEIPDYSSLSQAAANLAPP 419  
DB 352 FVPHLGSIQANGI---GSGNYGVLSWISPPSHPSGSGQVLDLWKIPNYGSSITEATHLAPS 408  
QY 420 VEPFGGREALVYFSAFPGPNRNSAPNDVPCLLPQEIYTHFVSEQAPTMGAALLHYDDP 479  
DB 409 VYPPGFGVLFVFMKMPGP---GAYNLPCLLPQEIYSHLASEQAPTVEAALLHYDDP 464  
QY 480 DTRNLGFEKFLYPGGLYTCVFNCGVAGAQOOLPLNGVFLFVSWRSFYQLKPVGTASTARS 539  
DB 465 DTGRNLGFEKAYPDGFLTCVFNCGVAGSGPQPLNGVFLFVSWRSFYQLKPVGTASTARS 524

QY 540 RLGVRR 545  
DB 525 RLGLRR 530

RESULT 4  
ADC72176  
ID ADC72176 standard; protein; 530 AA.  
XX ADC72176;  
XX ADC72176;  
DT 18-DEC-2003 (first entry)  
XX Norwalk virus protein 2 amino acid sequence.

immune response; non-Norwalk virus agent; immunogen; Norwalk virus;  
viral pathogen; acute gastroenteritis; virucidal; antiviral vaccine.

Norwalk virus.

US6572862-B1.

03-JUN-2003.

07-JUN-1995; 95US-00486049.

08-NOV-1989; 89US-00433492.

27-APR-1990; 90US-00515993.

27-AUG-1990; 90US-00573509.

06-MAY-1991; 91US-00696454.

(BAYU ) BAYLOR COLLEGE MEDICINE.

Estes MK, Jiang X, Graham DY;

WPI: 2003-776005/73.

N-PSDB; ADC72174.

Inducing an immune response against non-Norwalk virus agents, comprises  
administering an immunogen recombinantly expressed from a cDNA from  
Norwalk virus.

Example 4; SEQ ID NO 3; 45pp; English.

This invention relates to a novel method of inducing an immune response  
in an individual against Norwalk virus and non-Norwalk virus agents, by  
orally or parenterally administering an immunogen recombinantly expressed  
or synthesised from a cDNA of Norwalk virus given in the specification.  
Norwalk virus is one of the most important viral pathogens, causing acute  
gastroenteritis. The invention may be used for the development of  
compounds with virucidal activity or an antiviral vaccine. The present  
sequence is the amino acid sequence of a protein encoded by the Norwalk  
virus genome of the invention.

Sequence 530 AA;

Query Match 70.3%; Score 2034.5; DB 7; Length 530;  
Best Local Similarity 69.4%; Pred. No. 3.1e-178; Mismatches 17; Gaps 6;  
Matches 379; Conservative 69; Indels 17; Gaps 6;

QY 1 MMASKADPQSDAGSAGQGVPEVNTADPLPMEPVAGPTTAVATAGVNMIDPWVNNF 60  
DB 1 MMASKDATSSVDGASGAGQGVPEVNASDPLAMPVAGSSTAVATAGVNPDPWVNNF 60  
QY 61 VQSPGQETISPNNTPGDILFDLQGLPHNPLPFLSHLSQMYNGWGNMVRILLAGNAFSA 120  
DB 61 VQAPQGEFTISPNNTPGDVLFDLSGLPHNPLPFLSHLSQMYNGWGNMVRILLAGNAFTA 120  
QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPHVIAVTRTLEPIEMPLEDVNLYHTND-NQPTM 179  
DB 121 GKIIIVCCIPPPGGSHNLTIATLPPHVIADVRTLDPIEVLPLEDVNLYHTND-NQPTM 180

180 RLVCMLYTLPLRTGGSGNSDSFVWAGRVLTAPSSDFGLFVLPPTIEQKTRAFVTPNPL 239  
181 RLVCMLYTLPLRTGGGIG--DSFVWAGRVMTCPSPDFLFLVPTVEQKTRPFTLPL 238  
240 QTLNSRSPFSLIOGMILSPDASQVVFQNGRCLIDQLLGTTPATSGQLFVRGKINQGA 299  
239 SLSNSRAPLPISSMGISPDNVQSVQFQNGRCTLDRGLVGTTPVSLSHVAKIRGTSN--G 296  
300 RTLNLTEDVGKPFMAFDSPAPVFPDFGKCDWMHMRISKTPNNTSSGDPMRSVSVQTNVQ 359  
297 TVINLTDLGTTPFHPFEGPAPIGFPLDGGCDWHINMTQFGHSSQT-----QYDVTTPDT 351  
360 FVPHLGSIQDEVFENHPTDYIGTIEWISQSPPTPGTDINLWEIPDYGSSLSQAANLAPP 419  
352 FVPHLGSIQANGI---GSGNYGVLSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 408  
420 VPPGFGFGEALVYFVSAPFPNRRSAPNDVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDP 479  
409 VPPGFGFGEVLVFFMSKMPGP---GAYNLPCLLPQEIYSHLASEQAPTVGEAALLHYVDP 464  
480 DTNRNLGEFKLYPGGYLTCVPNGVAGPQOLPLNGVLFVSVWSRFPYQLKPVGTASTARS 539  
465 DTGRNLGEFKAYPDGFLTCVPNGASSGPQQLPINGVVFVSVWSRFPYQLKPVGTASSARG 524  
540 RLGVRR 545  
525 RLGLRR 530

RESULT 5  
AAB49701  
ID AAB49701 standard; protein; 530 AA.  
XX AAB49701;  
DT 04-APR-2001 (first entry)  
XX Small round structured virus protein SEQ ID 2.  
DE Small round structured virus; SRSV; food poisoning.  
KW Small round structured virus.  
OS Small round structured virus.  
XX WO200079280-A1.  
XX 28-DEC-2000.  
XX 22-JUN-2000; 2000WO-JP004095.  
XX 22-JUN-1999; 99JP-00175928.  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX (DENK-) DENKA SEIKEN KK.  
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX WPI; 2001-080848/09.  
XX N-PSDB; AAF29142.  
XX Kit for the detection and typing of small round-structured virus (SRSV)  
XX PT strains for investigation of food poisoning outbreaks, contains  
XX PT antibodies.  
XX Claim 1; Page 42-45; 84pp; Japanese.  
XX This invention relates to a kit for the detection and typing of small  
XX round structured virus (SRSV) strains. The kit contains antibodies  
XX directed against peptides represented in sequences AAB49700 - AAB49710,  
XX which are each SRSV strain specific. Polynucleotide sequences AAB20141 -  
XX AAB20151 represent cDNA encoding the strain specific proteins. The kit is  
XX used for detecting and typing strains of SRSV in order to prevent the  
XX spread of infection and to examine the epidemiology of outbreaks

SQ Sequence 530 AA;  
Query Match 70.1%; Score 2029.5; DB 4; Length 530;  
Best Local Similarity 69.4%; Pred. No. 8.9e-178;  
Matches 379; Conservative 68; Mismatches 82; Indels 17; Gaps 6;  
QY 1 MMASKADAPQSDAGSAGAGQLVPEVNTADPLPMEVAGPTTAVATAGQVNMIDKIVNNF 60  
Db 1 MMASKADATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSTAVATAGQVNPIDPWIINN 60  
QY 61 VQSPQGEFTISPNTPGDIILFDLQGLPHNLPSLHSONYNGWGNMVRVILLAGNAFSA 120  
Db 61 VQAPQGEFTISPNTPGDIILFDLQGLPHNLPSLHSONYNGWGNMVRVILLAGNAFSA 120  
QY 121 GKIIVCCVPPGFTSSSLTIAQAATLPHVIADVRLTEPLEMPLDVRNLYHTND-NQPTM 179  
Db 121 GKIIVCSIPPGFSGHNLTTAQAATLPHVIADVRLTEPLEMPLDVRNLYHTND-NQPTM 180  
QY 180 RLVCMLYTLPLRTGGSGNSDSFVWAGRVLTAPSSDFGLFVLPPTIEQKTRAFVTPNPL 239  
Db 181 RLVCMLYTLPLRTGGGIG--DSFVWAGRVMTCPSPDFLFLVPTVEQKTRPFTLPL 238  
QY 240 QTLNSRFPSLIOGMILSPDASQVVFQNGRCLIDQLLGTTPATSGQLFVRGKINQGA 299  
Db 239 SLSNSRAPLPISSMGISPDNVQSVQFQNGRCTLDRGLVGTTPVSLSHVAKIRGTSN--G 296  
QY 300 RTLNLTEDVGKPFMAFDSPAPVFPDFGKCDWMHMRISKTPNNTSSGDPMRSVSVQTNVQ 359  
Db 297 TVINLTDLGTTPFHPFEGPAPIGFPLDGGCDWHINMTQFGHSSQT-----QYDVTTPDT 351  
QY 360 FVPHLGSIQDEVFENHPTDYIGTIEWISQSPPTPGTDINLWEIPDYGSSLSQAANLAPP 419  
Db 352 FVPHLGSIQANGI---GSGNYGVLSWSPSPHSGSQVDLWKIPNYGSSITEATHLAPS 408  
QY 420 VPPGFGFGEALVYFVSAPFPNRRSAPNDVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDP 479  
Db 409 VPPGFGFGEVLVFFMSKIPGAYS---LPCLLPQEIYSHLASEQAPTVGEAALLHYVDP 464  
QY 480 DTNRNLGEFKLYPGGYLTCVPNGVAGPQOLPLNGVLFVSVWSRFPYQLKPVGTASTARS 539  
Db 465 DTGRNLGEFKAYPDGFLTCVPNGASSGPQQLPINGVVFVSVWSRFPYQLKPVGTASSARG 524  
QY 540 RLGVRR 545  
Db 525 RLGLRR 530

RESULT 6  
AAR57091  
ID AAR57091 standard; protein; 530 AA.  
XX AAR57091;  
XX 27-AUG-2003 (revised)  
XX 25-MAR-2003 (revised)  
XX 05-OCT-1994 (first entry)  
XX Small round virus SRSV/KY/89 capsid protein.  
DE pathogen; acute gastroenteritis; food poisoning; seafood contamination;  
XX diagnostic assay; human calcivirus; small round virus; SRSV; KY89;  
XX Norwalk virus; capsid protein.  
XX Small round structured virus.  
XX WO9405700-A2.  
XX 17-MAR-1994.  
XX 07-SEP-1993; 93WO-US008447.  
XX 07-SEP-1992; 92US-00941365.  
XX

PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
XX Matson DO, Estes MK, Jiang X, Graham DY;  
XX WPI; 1994-101125/12.  
XX N-PSDB; AAQ56832.  
XX DNA from Norwalk and related viruses - used for preparing prods. for use  
PT in diagnostic assays, detection and vaccines for Norwalk and related  
PT viruses.  
XX  
XX Example 7; Fig 13a; 156pp; English.  
XX  
XX The known sequence for Norwalk virus was used to obtain the sequence of  
CC other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool  
CC from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide  
CC cDNA sequence includes part of the polymerase region and the capsid  
CC region of the genome; the deduced amino acid sequences are AAR57092 and  
CC AAR57091, respectively. Expression of fragments and derivs. of Norwalk-  
CC related viruses permits development of diagnostic assays to detect  
CC antibodies, antigens, viral genetic material or antivirals. (Updated on  
CC 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS  
CC field.)  
XX  
XX Sequence 530 AA;  
SQ

Query Match 68.9%; Score 1993.5; DB 2; Length 530;  
Best Local Similarity 68.3%; Pred. No. 1.8e-174;  
Matches 373; Conservative 68; Mismatches 88; Indels 17; Gaps 6;  
QY 1 MMASKADAPQADGASGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNNIDPWVNF 60  
DB 1 MMASKADATSSVDGASVQLVPEVNASDPLAMDPAVAGSTAVATAGQVNNIDPWVNF 60  
QY 61 VOSQGETTISPNNTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMVRILLAGNAFSA 120  
DB 61 VOAPQGETTISPNNTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMVRILLAGNAFTA 120  
QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNLYHTND-NQPTM 179  
DB 121 GKIIIVCIPPGFGSQQLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNLYHTND-NQPTM 180  
QY 180 RLVCMLYTLPTLRTGGSGNSDSFVAVAGRVLTAPSSDFSLFLVPPTEIOKTRAFVNNIPL 239  
DB 181 RLVCMLYTLPTLRTGGSGTG--DSFVAVAGRVMTCPSPDFNLFVLPVTEOKTRPFTLPNLP 238  
QY 240 QTLNSRFPSTLQGMILSPDASOVVQFONGRCLIDQLLGTTPATSGQLFRVRGKINQGA 299  
DB 239 SLSLSRAPLPISGMIGISPDNVSQVQFONGRCLIDQLLGTTPATSGQLFRVRGKINQGA 296  
QY 300 RTLANLTVGDKGPFMAFDSFVAVAGRVLTAPSSDFSLFLVPPTEIOKTRAFVNNIPL 359  
DB 297 TVINLTLDGTFPHFEGFAPIGFDLGGCDWHINMTQFGHSSQI-----QYDVDTTPTD 351  
QY 360 FVPHLGSIOFDEVFNHPTGDIYGTIEWTSQSTPPTGTDINLWEIPDYGSSLSQAANLAPP 419  
DB 352 SVPHLGSIOANGI--GSGNYIGVLSVSPSPHPSGQVDLWKIPNYGSSITEATHLAPS 408  
QY 420 VFPFGFGEALYFVSAPFPGNNRSAPNDVPCLLPOEYITHFVSEQAPTMGDAALHYVDP 479  
DB 409 VYSPFGFGEVLFVFSKIPFGGDS-----LPCLLPQGYISHLASEQAPTVGEGPLHYVDP 464  
QY 480 DTRNLGFEKLYPGGGLYLCVNGVAGPQQQLPLNGVFLFVSVRSFYQLKPVGTASTARS 539  
DB 465 DTRNLGFEKAYPDGFLTCVNGAGSSGQQQLPFGVFFVSVRSFYQLKPVGTASTARG 524  
QY 540 RLGVRR 545  
DB 525 RLGLRR 530

RESULT 7  
AAB49700

ID AAB49700 standard; protein; 545 AA.  
XX  
AC AAB49700;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 1.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX  
PA (NTNA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
XX  
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX  
DR WPI; 2001-080848/09.  
DR N-PSDB; AAF29141.  
XX  
PT Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX  
PS Claim 1; Page 40-42; 84pp; Japanese.  
XX  
CC This invention relates to a kit for the detection and typing of small  
CC round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX  
SQ Sequence 545 AA;  
Query Match 68.3%; Score 1977.5; DB 4; Length 545;  
Best Local Similarity 66.6%; Pred. No. 5.8e-173;  
Matches 367; Conservative 76; Mismatches 95; Indels 13; Gaps 7;  
QY 1 MMASKADAPQADGASGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNNIDPWVNF 60  
DB 1 MMASKADAPTNDGTSGAGQLVPEANTAEPISEMPEVAGATAAATAGQVNNIDPWVNF 60  
QY 61 VOSQGETTISPNNTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMVRILLAGNAFSA 120  
DB 61 VOAPQGETTISPNNTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMVRILLAGNAFTA 120  
QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNLYHTNDNQPTMR 180  
DB 121 GKIIIVCIPPGFQAQNTSIAQATLPPHVIADVRLTLEPIEMPLEDVRNLYHTNDNQPTMR 180  
QY 181 LVCMLYTLPTLRTGGSGNSDSFVAVAGRVLTAPSSDFSLFLVPPTEIOKTRAFVNNIPLQ 240  
DB 181 LVCMLYTLPLRASGSSGTDPPVIAGRVLTCPSPDFSLFLVPPNVEQTKPFSVPLNPLN 240  
QY 241 TLSNRFPSTLQGMILSPDASOVVQFONGRCLIDQLLGTTPATSGQLFRVRGKI--NOG 298  
DB 241 TLSNRFPSTLKSMMVSRDHGQVQFONGRVTLTQGLQGTTPTSASQJCKIRGSVFHANG 300  
QY 299 ARTLNLTVEGDKPFMAFDSFVAVAGRVLTAPSSDFSLFLVPPTEIOKTRAFVNNIPLQ 358  
DB 301 GNGYNLTGDSYPHAFESPAIGFPDGLGCDWHMEASPT-TQFNTGDIYKQINVKQES-S 358  
QY 359 GFVPHLGSIOFDEVFNHPTG-DYIGTIEWISQSTPPTGTDINLWEIPDYGSSLSQAANLA 417

Wed Jun 2 09:13:27 2004

us-09-926-799-3.rag

```

Db 359 AFAPHLGTIQADCLSDVSNTNMIAKLGWSPVSDGHRGDVDPVPIPRYGSTLLEAAQLA 418
Qy 418 PVFPFPGFGEALVYFVSAPP---GNNSAPNDVPCLLPOEYITHFYSEQAPTWGDAALL 474
Db 419 PPIYPGFGFGEATVFFMSDFPIAHGTNGLS---VPCITPQBFVTHFVNEQAPTGEAALL 474
Qy 475 HVVDPTNRNLGEEFKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWVSRYQLKPVGTA 534
Db 475 HYLDDPTNRNLGEEFKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWVSRYQLKPVGTA 534
Qy 535 STARSRLGVRR 545
Db 535 GPA-CRLGIRR 544

RESULT 8
AAB49706
ID AAB49707 standard; protein; 535 AA.
XX
AC AAB49707;
DT 04-APR-2001 (first entry)
DE Small round structured virus protein SEQ ID 8.
KW Small round structured virus; SRSV; food poisoning.
XX Small round structured virus.
XX WO200079280-A1.
XX 28-DEC-2000.
XX 22-JUN-2000; 2000WO-JP004095.
XX 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
DR N-PSDB; AAF29148.
XX
PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
PS Claim 1; Page 57-59; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
SQ Sequence 535 AA;
Query Match 40.6%; Score 1175; DB 4; Length 535;
Best Local Similarity 46.2%; Pred. No. 5.2e-99;
Matches 261; Conservative 82; Mismatches 172; Indels 50; Gaps 16;
Qy 1 MMWAKDAPOSADGASGAGQLVPEVNTADPLPMEPVAGETTAATAGQVNNIDPWINVF 60
Db 1 MKWASNDAPSNDGAG---LVPEANN-ETMALEPVAGASIAAPLTGQNNIIDPWIRLNF 56
Qy 61 VQSPQGEFTSPNTPGDIPLDLQGLPHLNPFLSHLSOMYNGVGNMVRILLAGNATSA 120
Db 57 VOAPNGEFTVSPRNSPGEVLNLELGLFELNPLAHLRSRYINGIAGGVEVQVILLAGNATFA 116

```

```

Qy 121 GKIIVCCVPBGTSSTLTIAQATLPHVITADVRLTEPIEMPLEDVRNLYHTN-DNQPTM 179
Db 117 GKLVFAAVPFPHPLENIPEGQITMFPHTIIDVRLTEPLVLLPDPVRNNFFHYNQNEPRM 176
Qy 180 RLVMCLYTLRTTGGSGNSDSFVWAGRVLTAPSSDFSFLFVPPTEIQKTRAFVNIPL 239
Db 177 RLVMALYTLPLRS-NGSGD-DVFTVSCRVLTRPSDPDFENLYLPPTLESKTKPTTLPILTI 234
Qy 240 QTLNSRFRPSLIQCMTLSPDASQVQFONGRCIDGQLLGTTPATSGQLFRVRGKINQGA 299
Db 235 GELTNSRFPVIDELYTSPNESLVQPNQRCALDGELOQTTLQLLPTAICSFGRINQKV 294
Qy 300 R-----TLNLTEVDGKPF-MAFDSAPVGFDPF-CKCDMHRIRISKTPNNTSSGDPMSRV 351
Db 295 SGENHWNMQVTNINGTPDPTGDPAPLGTDFSGKLFGLVLSQRDHDNACRSHDAV--- 351
Qy 352 SVQTNVQGFVPHLGSIOFDEVFNHPTGDYIGTIE----WISOPS---TPPG---TDINLW 401
Db 352 -IATNSAKFTPKLGAIQ-----IGTWEEDDVHINQPTKFTPVGLFENEGFNQW 398
Qy 402 EIPDYGSSLSQAANLAPPVFPFGFGEALVYFVSAPFPNNRSAPNDVPCLLPOEYITHFV 461
Db 399 TLPNYSGALTNNMGLAPPVAPTFPGEQILFFRSHIPLKGGVADP-VIDCLLPQEWIQLHY 457
Qy 462 SEQAPTMGDAALLHYVDPTNRNLGEPKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSW 521
Db 458 QESAPSQSDVALIRFTNPTDGRVLFEAKLRSGYITVA--NTGSRPIVVPANGFRFDTIW 515
Qy 522 VSRFYQLKPVGTASTARSRLGVRR 546
Db 516 VNQFYSLAPMGTGN-----GRRKV 534

RESULT 9
AAB49706
ID AAB49706 standard; protein; 540 AA.
XX
AC AAB49706;
DT 04-APR-2001 (first entry)
DE Small round structured virus protein SEQ ID 7.
KW Small round structured virus; SRSV; food poisoning.
XX Small round structured virus.
XX WO200079280-A1.
XX 28-DEC-2000.
XX 22-JUN-2000; 2000WO-JP004095.
XX 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
DR N-PSDB; AAF29147.
XX
PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
PS Claim 1; Page 54-57; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,

```



CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX  
 SQ Sequence 540 AA;

Query Match 40.4%; Score 1169.5; DB 4; Length 540;  
 Best Local Similarity 44.9%; Pred. No. 1.7e-98;  
 Matches 254; Conservative 87; Mismatches 178; Indels 47; Gaps 13;  
 QY 1 MMASKDAPQSDAGSAGAGQLVPEVTADPLMEPVAGPTTAVATAGVNMIDPMVNF 60  
 DB 1 MMASNDATPSNDGAG--LVPSNN-EAMALEPVVGASLAAPVTGQNIIDPIRNF 56  
 QY 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSOMYNGWGNMRVILLAGNAFSA 120  
 DB 57 VQAPNGEFTVSPRNSPGEILNLELGPENLPYLAHARMYNGYAGMEVQVNLGNAFTA 116  
 QY 121 KGIIVCCVPPGPTSSSLTIAQATLPHVIAVRLTLEPIEMPLEDVRNVLHTND-NQPTM 179  
 DB 117 KGIIFAAVPPVPPVENLSPSQITMPPHVIDVRLTLEPVLPMFVDRSTLHFHFNQDEPKM 176  
 QY 180 RLVCMLYTLRTGGSGNSDSFVAGRVLTAPSSDFSLFLVPTTIEOKTRAFVTPNIP 239  
 DB 177 RLVALMYTPLRS-NGSGD-DVFTVSCRILTRESPEFDFTYLPPTVESKTPFTLPVLT 234  
 QY 240 QTLNSRFPPLTQGMILSPDASOVQFONGRCLIDGLLGTTPATSGOLFRVRGKI---- 295  
 DB 235 GELNSRFPPLSIDEMVTSFNSIVVQPNQGRVTLDELGLLTQLOACNICSRIGKVTQV 294  
 QY 296 --NQGARTLNLTVEVGKPFMAFDS-PAPVGFDPFGKCDWMHRISKTPNNTSSGDPMSVS 352  
 DB 295 PSEOHMNLITNLNGTQDPTDDVPALGVDFAGEVGLVSQRNGESNPANRAHDAV 354  
 QY 353 VQTNVQGVPHLGSIQF-----DEVFNHTGDIYGTIEWISOPSTPGTD-----INL 400  
 DB 355 VATVSDKTPKGLVQIGTWNNDVENQPT-----KFTPIGLNEVANGHRFEQ 402  
 QY 401 WEIPDYGSSLSQAANLAPPVPPGFGELVYFVSAPPGNRNSAPNDVPCLLPQBYIYTHF 460  
 DB 403 WTLFRYSGALTINMLAPAVAPLPTGERLLLFRRSVVPLKGGFNP-AIDCSVPQEWQHF 461  
 QY 461 VSEQAPTMGDAALHYVDPDTRNRLGEPKLYPGGYLTCPVNGVAGPQOLPLNGVLFVS 520  
 DB 462 VQESAPSLGDVALRVYNPDTRGLVFEAKLHKGGLTV--SSTSTGTPVVPANGYFKFDS 519  
 QY 521 WVSRYQLKPVGTASTARSRLGVRI 546  
 DB 520 WVNQFYSLAPMGTCN-----GRRRV 539

RESULT 10  
 AAB49708  
 ID AAB49708 standard; protein; 542 AA.  
 XX  
 AC AAB49708;  
 DT 04-APR-2001 (first entry)  
 DE Small round structured virus protein SEQ ID 9.  
 XX  
 KW Small round structured virus; SRSV; food poisoning.  
 XX  
 OS Small round structured virus.  
 XX  
 PN WO200079280-A1.  
 XX  
 PD 28-DEC-2000.  
 PF 22-JUN-2000; 2000WO-JP004095.  
 XX  
 PR 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 PA (DENK-) DENKA SEIKEN KK.  
 XX  
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 XX  
 DR WPI; 2001-080848/09.  
 DR N-PSDB; AAF29149.  
 XX  
 PT Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.  
 XX  
 PS Claim 1; Page 59-61; 84pp; Japanese.  
 XX  
 CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX  
 SQ Sequence 542 AA;  
 Query Match 39.9%; Score 1155.5; DB 4; Length 542;  
 Best Local Similarity 45.3%; Pred. No. 3.4e-97;  
 Matches 258; Conservative 88; Mismatches 171; Indels 53; Gaps 18;  
 QY 1 MMASKDAPQSDAGSAGAGQLVPE-VNTADPLMEPVAGPTTAVATAGVNMIDPMVNF 59  
 DB 1 MMASNDAPNDGAA---SLVPEGIN--ETMPLEPVAGASIAAPVAGQNTIIDPIRNF 55  
 QY 60 FVQSPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSOMYNGWGNMRVILLAGNAFS 119  
 DB 56 FVQAPNGEFTVSPRNSPGEILNLELGPENLPYLAHARMYNGYAGMEVQVNLGNAFTA 115  
 QY 120 AKIIVCCVPPGPTSSSLTIAQATLPHVIAVRLTLEPIEMPLEDVRNVLHTNDN-QPT 178  
 DB 116 AGKILFAAIPENFLVDMISPAQITMLPHLIVDVRTLEPITPLDVRNVFYHFNQPOPR 175  
 QY 179 MRLVCMLYTLRTGGSGNSDSFVAGRVLTAPSSDFSLFLVPTTIEOKTRAFVTPNIP 238  
 DB 176 MRLVALMYTPLRS-NGSGD-DVFTVSCRILTRESPEFDFTYLPPTVESKTPFTLPILT 233  
 QY 239 LOTLSNSRFPPLTQGMILSPDASOVQFONGRCLIDGLLGTTPATSGOLPRVRGKI--- 295  
 DB 234 ISELNSRFPPIEOLYTAPNETNVVQCNGRCLIDGELQGTTLSSAVCFLOQRTVAD 293  
 QY 296 --NQGARTLNLTVEVGKPFMAFDS-PAPVGFDPFGKCDWMHRISKTPNNTSSGDP--MR 349  
 DB 294 NGDNWDQNLQLTYPENGASYDPTDEVPAPLGTQDFSGMLYGV-LTQDNVNVSTGEAKNAK 352  
 QY 350 SVSVQTNVQGVPHLGSIQFDEVFN--HPTGDYGTIEWISQPS--TPPG-----TDIN 399  
 DB 353 GIYISTTSKFTPKIGSIGLSIETHVHP-----NQQRFTPVGVAVENTPQF 401  
 QY 400 LWEIPDYGSSLSQAANLAPPVPPGFGELVYFVSAPP--GPNRNSAPNDVPCLLPQBY 456  
 DB 402 QMVLPHYAGSLALNTNLAPAVAPTTPGQQLFFRSRVPVCOLOQODA--FIDCLLPQEW 459  
 QY 457 ITHFVSEQAPTMGDAALHYVDPDTRNRLGEPKLYPGGYLTCPVNGVAGPQOLPLNGVVF 516  
 DB 460 VNHFYQEAAPQADVALRVYNPDTRGLVFEAKLHRSGFITV--SHTGAYPLVVPNGHF 517  
 QY 517 LFGVNSRFFYQLKPVGTASTARSRLGVRI 546  
 DB 518 RFDSWNVQFYSLAPMGTCN-----GRRRI 541

RESULT 11  
 AAB49705  
 ID AAB49705 standard; protein; 548 AA.

351 RNLSTTRAEAKVDTTAGRTFKLGSLEISTSDDEFDQ--NQPT-----KFT 396  
393 PFG-----TDINLWEIPDYSSLSQAANLAPPYPPGFGGALVYFVSAPFPNNRSAPN 446  
397 PVGIGVDNEABFQWLSLDPYSGQFTNNMLAPAVAFNPFQQLFFRSQLSGGRS--N 454  
447 DV-PCLLPOEYITHFVSEQATMGDAALLHYVDPDTNRNLGFKLYPGGYLTCVPGVGA 505  
455 GVLDCLVPQEWQHFQYESAPAQTOVALVRYVNDTGKVLPEAKHLKLGFTIANN--GD 512  
506 GPOQLPLNGVFLFVSVSRFYQLKPVGTASTARSRLGVRR 546  
513 SPITVPNGVFRFESWVNDFTLAPMTGN-----GRRRI 547

RESULT 12  
AAU91272  
ID AAU91272 standard; protein; 548 AA.

XX AC AAU91272;  
XX DT 18-JUN-2002 (first entry)  
XX DE Norwalk virus associated polynucleotide #1.  
XX DE Norwalk virus; monoclonal antibody; geno group I; geno group II;  
XX KW immunological detection; food; viral infection.  
XX OS Norwalk virus.  
XX PN JP2002020399-A.  
XX PD 23-JAN-2002.  
XX PF 10-JUL-2000; 2000JP-00208151.  
XX PR 10-JUL-2000; 2000JP-00208151.  
XX PA (OSAP ) OSAKA PREFECTURE.  
XX PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.  
XX PA (IATR ) IATRON LAB INC.  
XX WPI; 2002-287412/33.

XX A monoclonal antibody useful in the immunological detection and diagnosis  
XX of Norwalk virus infection.  
XX PS Disclosure; Page 12-13; 24pp; Japanese.

XX CC The invention describes a monoclonal antibody recognising Norwalk virus,  
XX a capsid protein of Norwalk virus, or a common antigen epitope on the  
XX capsid protein molecule of geno group I and geno group II. The antibody  
XX is useful for immunological detection and quantitative analysis of  
XX Norwalk virus in foods and the serum of infected patients. This sequence  
XX represents a Norwalk virus associated protein described in the invention  
XX  
XX SQ Sequence 548 AA;

Query Match 38.5%; Score 1124.5; DB 4; Length 548;  
Best Local Similarity 43.0%; Pred. No. 2.5e-94; Indels 69; Gaps 17;  
Matches 250; Conservative 80; Mismatches 182;

QY 1 MMASKDAPOQADGASGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWVNNF 60  
DB 1 MKMASNDAPNDGAG--LVPEINN-EAMALDPVAGAAIAAPLTGQNIIDPWIMNNF 56  
QY 61 VQSPQGEFTISPNNTPGDILFDLQGLPHNLPFLSHLSQMYNGWGNVRVILLAGNAFSA 120  
DB 57 VQAPGGEFTVSPRNSPGEVLLNLELGPENPYLAHARMYNGYAGGFEVQVVLGNAFTA 116  
QY 121 GKIIIVCCVPPGFTSSSLITTAQATLFPVHIADVRLTEPLEMPLDVRNVLYHTNDNQPT-M 179  
DB 117 GKIIIFAAIPPNPFDNLSSAAQITMCPHVIVDVRQLEPNLPMFVVRNFFHYNQSDSRL 176

XX AC AAB49705;  
XX DT 04-APR-2001 (first entry)  
XX DE Small round structured virus protein SEQ ID 6.  
XX KW Small round structured virus; SRSV; food poisoning.  
XX OS Small round structured virus.  
XX PN WO200079280-A1.  
XX PD 28-DEC-2000.  
XX PF 22-JUN-2000; 2000WO-JP004095.  
XX PR 22-JUN-1999; 99JP-00175928.  
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX PA (DENK-) DENKA SEIKEN KK.  
XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX DR WPI; 2001-080848/09.  
XX DR N-PSDB; AAF29146.  
XX PT Kit for the detection and typing of small round-structured virus (SRSV)  
XX PT strains for investigation of food poisoning outbreaks, contains  
XX PT antibodies.  
XX PS Claim 1; Page 52-54; 84pp; Japanese.

XX CC This invention relates to a kit for the detection and typing of small  
XX round structured virus (SRSV) strains. The kit contains antibodies  
XX directed against peptides represented in sequences AAB49700 - AAB49710,  
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
XX used for detecting and typing strains of SRSV in order to prevent the  
XX spread of infection and to examine the epidemiology of outbreaks  
XX  
XX SQ Sequence 548 AA;

Query Match 38.8%; Score 1124.5; DB 4; Length 548;  
Best Local Similarity 43.0%; Pred. No. 2.5e-94; Indels 69; Gaps 17;  
Matches 250; Conservative 80; Mismatches 182;

QY 1 MMASKDAPOQADGASGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWVNNF 60  
DB 1 MKMASNDAPNDGAG--LVPEINN-EAMALDPVAGAAIAAPLTGQNIIDPWIMNNF 56  
QY 61 VQSPQGEFTISPNNTPGDILFDLQGLPHNLPFLSHLSQMYNGWGNVRVILLAGNAFSA 120  
DB 57 VQAPGGEFTVSPRNSPGEVLLNLELGPENPYLAHARMYNGYAGGFEVQVVLGNAFTA 116  
QY 121 GKIIIVCCVPPGFTSSSLITTAQATLFPVHIADVRLTEPLEMPLDVRNVLYHTNDNQPT-M 179  
DB 117 GKIIIFAAIPPNPFDNLSSAAQITMCPHVIVDVRQLEPNLPMFVVRNFFHYNQSDSRL 176  
QY 180 RLVCMLYPLRTGGSGNSDSFVAVAGRLTAPSSDFSFLVPPTEIOKTRAFVNIPL 239  
DB 177 RLIIAMLYPLR--ANNSGDVFTVSCRVLTRPSDFSNFLVPPTEVSKTKPFTLPILT 234  
QY 240 QTLNSRPFSLIQGMILSPASQVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINGOA 299  
DB 235 SEMSNRFPVPIESLRTENIVVQCONGRVTLDELMTGLQLLPSQICAFRGVLRST 294  
QY 300 -----RTLN-----LTEVDGKPF-MAFDSAPVGFDPF-GKCDWHMRISKTP 339  
DB 295 SRASDAQDTATPFLFNYWHVQLDNLNGTFYPDAEDIPGLTPTDFRGK---VFGVASQ 350  
QY 340 NNTSSGDMRVSQVNVQGFPHLSIQ-----FDEVFNHPTGDIYGTIEMISQPS 392



Wed Jun 2 09:13:27 2004

us-09-926-799-3.rag

```

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
XX DR N-PSDB; AAF29150.
XX PT Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX PS Claim 1; Page 62-64; 84pp; Japanese.
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX CC
XX CC Sequence 550 AA;
XX CC
XX CC Query Match 37.7%; Score 1090.5; DB 4; Length 550;
XX CC Best Local Similarity 42.8%; Pred. No. 3.4e-91;
XX CC Matches 250; Conservative 75; Mismatches 173; Indels 89; Gaps 16;
XX CC
QY 1 MMASKDAPQADGASGAGQVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWINNF 60
DB 1 MKWASNDAPSNDGAA---NLVPEAND-EVMALEPVVGASIAAPVVGQNIIDPWIENF 56
QY 61 VQSPQGEFTISNNTPGDIILFDLQGLPHLNPFLSHLSQMYNGVGNMVRILLAGNAPSA 120
DB 57 VQAPQGEFTVSPRNSPGEMLNLELGPDLNPNYLAHARMYNGHAGMEVQIVLAGNAFTA 116
QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPHPHVIAVTRLEPIEMPLEDVRNVLVHTND-NOPTM 179
DB 117 GKIIIFAAVPPHVFENISAAQITMCPHVIDVRQLEPVLLPLDIRNRPFFHYNQENTPRM 176
QY 180 RLVCMLYTPRLTGGSGNSDSFVVAGRVLTAPSSDFSFLVPPTEQKTRAFVNPIL 239
DB 177 RLVMALYTPLRANSQ---EDVFTVSCRVLTRPAPDEFTELVPPTVESKTKPFTLPILTL 233
QY 240 QTLNSRFPSSLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRGK-INQG 298
DB 234 GELNSRFPAAIDMLYTDPNESIVVQFQNGRCLIDGLTQGTQTLVPTQICAFRGTLISQT 293
QY 299 ARTNLNTE-----VDGKPFMAFDS-PAPVGFDPDFGKCDWHMRISKTPNN 341
DB 294 ARAADSTDSQPARNHPLHVQVKNLDGTYDPTDDIPAVLGAIDF-----FVPHLSIQF 376
QY 342 TSSGDPMRSVSVQTNVQ-----FVPHLSIQF-----DEVFNHP 376
DB 341 TVFG-----VASORDVQSQEQGHYATRAHEAHDITDDPKYAPKGLTILIKSGSDDFN-- 393
QY 377 TCDYIGTIEWISQP--STPPGTDIN---LWEIPDYGSSLSQAANLAPPVPPPGFEALVY 431
DB 394 -----THQPIRFTFVGGNDNNWQWELPDYSGRLTLNMLNLAAPSFPGERILF 443
QY 432 FVSAPFPGNRSAPNDPCLLPQEVYTHFVSEQAPTMGDAALLHYVDPDTNRLGEEKLY 491
DB 444 FRSIVPSAGGYGS-GYIDCLIPQEWVQHFYQEAAPSQAVALRVYVNPDTGRNIFEAKHL 502
QY 492 PGGYLTCVPGVAGCPQQLPLNGVFLFVSWVSREYOLKPVGTASTAR 538
DB 503 REGFLTVA--NCGNNPIVPPNGYFRFEAWGNQFTLAPMSGQGR 547
XX CC
XX CC RESULT 15
XX CC ID AAB49710
XX CC AAB49710 standard; protein; 541 AA.
XX AC
XX AAB49710;
XX DT 04-APR-2001 (first entry)

```

Db 412 T-----SPTTP--TWQOWRCSRGNCCSSGHRYPVPMNRVTWIVLSHKSGFSTSTR 464  
QY 442 RSAPNDVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDPDTNRNLGFEKLYPGGYLTCVEN 501  
Db 465 K-----LPQ-----LNLRW-----LIRFINPDTGRVLPFARLHKQGFITVA-- 501  
QY 502 GVGAGPQQLPLNGVFLFVSWSEFYQLKPVGTASTARSLGVRI 546  
Db 502 HTGDNPVMPNGYPRFEAWNQFYSLAFVGTGK-----GRRRV 540

Search completed: June 1, 2004, 13:45:48  
Job time : 48.9047 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 13.0087 Seconds  
(without alignments)  
2166.837 Million cell updates/sec

Title: US-09-926-799-3  
Perfect score: 2895  
Sequence: 1 MMASKADPQSGADGASGAGQ.....QLKPVGTASTARSRLGVRR1 546

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034.5	70.3	530	US-08-486-049-3	Sequence 3, Appli
2	290	10.0	669	US-09-617-594A-2	Sequence 2, Appli
3	272	9.4	668	US-09-617-594A-4	Sequence 4, Appli
4	268.5	9.3	626	US-09-590-020-7	Sequence 7, Appli
5	267	9.2	623	US-09-590-020-2	Sequence 2, Appli
6	267	9.2	623	US-09-590-020-4	Sequence 4, Appli
7	256	8.8	622	US-09-590-020-6	Sequence 6, Appli
8	140	4.8	2206	US-07-852-260-2	Sequence 2, Appli
9	140	4.8	2206	US-08-461-503-2	Sequence 2, Appli
10	140	4.8	2206	US-08-465-250-2	Sequence 2, Appli
11	123	4.2	2972	US-09-579-181-2	Sequence 2, Appli
12	123	4.2	3118	US-09-579-181-1	Sequence 1, Appli
13	120	4.1	3443	US-08-416-603-2	Sequence 2, Appli
14	109.5	3.8	2736	US-09-252-991A-30227	Sequence 30227, A
15	109	3.8	2322	US-09-976-594-15	Sequence 15, Appli
16	107	3.7	1048	US-09-171-699-10	Sequence 10, Appli
17	106	3.7	2318	US-09-091-219-24	Sequence 24, Appli
18	106	3.7	2318	US-09-660-541-24	Sequence 24, Appli
19	105.5	3.6	580	US-08-672-564-1	Sequence 1, Appli
20	102.5	3.5	1911	US-09-854-856-64	Sequence 64, Appli
21	102.5	3.5	1939	US-09-854-856-48	Sequence 48, Appli
22	102.5	3.5	1971	US-09-854-856-32	Sequence 32, Appli
23	102.5	3.5	1999	US-09-854-856-16	Sequence 16, Appli
24	102.5	3.5	2004	US-09-854-856-58	Sequence 58, Appli
25	102.5	3.5	2032	US-09-854-856-42	Sequence 42, Appli
26	102.5	3.5	2048	US-09-854-856-62	Sequence 62, Appli
27	102.5	3.5	2064	US-09-854-856-26	Sequence 26, Appli

28	102.5	3.5	2076	4	US-09-854-856-46	Sequence 46, Appli
29	102.5	3.5	2092	4	US-09-854-856-10	Sequence 10, Appli
30	102.5	3.5	2108	4	US-09-854-856-30	Sequence 30, Appli
31	102.5	3.5	2136	4	US-09-854-856-14	Sequence 14, Appli
32	102.5	3.5	2141	4	US-09-854-856-56	Sequence 56, Appli
33	102.5	3.5	2157	4	US-09-854-856-52	Sequence 52, Appli
34	102.5	3.5	2169	4	US-09-854-856-40	Sequence 40, Appli
35	102.5	3.5	2185	4	US-09-854-856-36	Sequence 36, Appli
36	102.5	3.5	2201	4	US-09-854-856-24	Sequence 24, Appli
37	102.5	3.5	2217	4	US-09-854-856-20	Sequence 20, Appli
38	102.5	3.5	2229	4	US-09-854-856-8	Sequence 8, Appli
39	102.5	3.5	2245	4	US-09-854-856-4	Sequence 4, Appli
40	102.5	3.5	2294	4	US-09-854-856-50	Sequence 50, Appli
41	102.5	3.5	2322	4	US-09-854-856-34	Sequence 34, Appli
42	102.5	3.5	2354	4	US-09-854-856-18	Sequence 18, Appli
43	102.5	3.5	2382	4	US-09-854-856-2	Sequence 2, Appli
44	100.5	3.5	580	4	US-09-252-991A-23699	Sequence 23699, A
45	100.5	3.5	721	4	US-09-390-234-20	Sequence 20, Appli

ALIGNMENTS

RESULT 1  
US-08-486-049-3  
; Sequence 3, Application US/08486049  
; Patent No. 6572862  
; GENERAL INFORMATION:  
; APPLICANT: BEIG, Xiang, Xi  
; APPLICANT: Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Fulbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE:  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-486-049-3

Query Match 70.3%; Score 2034.5; DB 4; Length 530;  
Best Local Similarity 69.4%; Pred. No. 1.9e-193;  
Matches 379; Conservative 69; Mismatches 81; Indels 17; Gaps 6;  
QY 1 MMASKADPQSGADGASGAGQLVPEVNTADPLPMEVAGPTTAVATAGQVNDPVIINNF 60  
DB 1 MMASKDATSSVDGASGAGQLVPEYNASDPLAMDPLVAGSTAVATAGQVNDPVIINNF 60

QY 61 VSPQGEFTISNNTPGDILOLQPHNLPFLSHLSQMYNGWGNMVRILLAGNAFSA 120  
Db 61 VQAPQGEFTISNNTPGDILOLQPHNLPFLSHLSQMYNGWGNMVRILLAGNAFTA 120  
QY 121 GKIIIVCCVPPGPTSSSLTIAQATLFPFHVIADVRTLEPIEMPLEDVRNVLVYHTND-NQPTM 179  
Db 121 GKIIIVCICPPGGSNHLTIAQATLFPFHVIADVRTLEPIEMPLEDVRNVLVYHTND-NQPTM 180  
QY 180 RLVCMLYTPRTGGSGNSDSFVAVGRVLTAPSSDFSLFVPPPTIEOKTRAFVTPNIPL 239  
Db 181 RLVCMLYTPRTGGGTG--DSFVAVGRVMTCPSPDFNLFVPPVTEOKTRFPFLPNLPL 238  
QY 240 OTLSNRFSLIOGMLTSPDASOVQFQNGRCLIDQLLGTTPATSGOLFRVRGKINOGA 299  
Db 239 SSLSNRAPLPTSSMIGISPDNVQSVQFQNGRCLTDLGRVLTTPVSLSHVAKIRGTSN--G 296  
QY 300 RTNLNTEVDGKPPMAFDSAPVGFDFGKCDWHMRISKTPNNTSSGDMRSPSVQTNVQG 359  
Db 297 TWINLTLDGTPHPPEGPAPIGFDPDLGGCDWHINNTQFHSQT-----QYDVTTPDT 351  
QY 360 FVPHLSIQDFEVNHTPDYICTIEWISOPSTPPGTDINLWEIPDYGSSLSQAANLAPP 419  
Db 352 FVPHLSIQANGI---GSGNYGVLSWISPPSHPSGSQVDLWKIPNYGSSITEATHLAPS 408  
QY 420 VFPPGFGELVYFVSAPFPPNRSAPNDVPCLLPQBYIYTHFVSEQAPTMGDAALLHYVDP 479  
Db 409 VYPPGFGELVYFVSMKMPG---GAYNLPCLLPQBYIYSHLASEQAPTMGEAALLHYVDP 464  
QY 480 DTRNLGEFLKYPGGYLTCPVNGVAGPQOLPLNGVFLFVSWSRFYQLKPVGTASTARS 539  
Db 465 DTGRNLGEFKAYPDGFLTCVPNGASSGQQLPINGVVFVSWSRFYQLKPVGTASSARG 524  
QY 540 RLGVRR 545  
Db 525 RLGVRR 530

RESULT 2  
US-09-617-594A-2  
; Sequence 2, Application US/09617594A  
; Patent No. 6541458  
; GENERAL INFORMATION:  
; APPLICANT: Audomnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.1  
; CURRENT APPLICATION NUMBER: US/09/617,594A  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 669  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-617-594A-2

Query Match 10.0%; Score 290; DB 4; Length 669;  
Best Local Similarity 25.5%; Pred. No. 9.8e-20;  
Matches 153; Conservative 71; Mismatches 247; Indels 130; Gaps 29;

QY 12 ADGASGAGOLVPEVNT-ADPLPMPVAGPTTAVATAGQVNMIDPWINVFNVSQGEFTI 70  
Db 125 ADGDSII--TTPEQGLVGVIAEPAQMAATAADATGKSDSEW--ESFF-SFITSVNW 179  
QY 71 SPNNTPGDILFDLQPHNLPFLSHLSQMYNGWGNMVRILLAGNAFSAKIIIVCCVPP 130  
Db 180 STSETQKILFKQSLGPLNLPFLSHLSKLYVAVSGSVDFRFSISGSGVGGKLAIVVPP 239

QY 131 GFTS-SSLTIAQATLFPFHVIADVRTLEPIEMPLEDVRNVLVYHTNDNQPTMELVCMLYTPL 189  
Db 240 GVDVFOVSTMLQ--YPHVLFARQVEPVIPIPIPLRSTLYHLMSDQTTSLIVIMVYNDL 296  
QY 190 -RTGGSGNSDSFVAVGRVLTAPSSDFSLFVPPPTIEOKTRAFVTPNIPLQTLNS--- 245  
Db 297 INPYANDSNSSGCIVT--VETKPGDPFKPHLLKPP-----GSMLTGHSIPSLIPKSSSL 349  
QY 246 ----RPPSLIOGMILSPDASOVVQNGRCLIDQLLGTTPATSGOLFRVRGKINOGART 301  
Db 350 WIGNRYMSDITDFVIRP-----FVFOANRHFDFNQ--ETAGWSTPRFR-----PIT 393  
QY 302 LNLTEVDGKPF--MAFDSAPVGFDFGKCDWHMRISKTP-----NNTSSGDP 347  
Db 394 ITLSNGSKLGTGVATDVIIP-GIPD-GWPDITTEELTPAGDYSITNGSGNDIATANA 451  
QY 348 MRSVSVQTNVQGVPHLGSIQDFEVNHTPDYI-CTIB--W-----ISOP 390  
Db 452 YDSADVITNTNF-----RGMVICALQRAMGDKKISSTAITTAIKG 495  
QY 391 ST-PPGTDINLWEIPDY---GSSLSQAANLAPPVFPFPGFGEAL-----VYFVSAPP 437  
Db 496 NTLKPSNTIDMTKIAVQDTHVGRDVTSDDTLAILGYTGIGQAIGSNRDSVVRISMLP 555  
QY 438 GPNNRSA-----PNDVPCLLPQBYIYTHFVSEQAPTMGDAALLHYVDP-----DT 481  
Db 556 ETGARGGNHPFYKNSIKGLVLRSDIVFNSQILHTSRQSLNHYLLPPDPSFAVYRIIS 615  
QY 482 NRNLGEFLKYPGGYLTCPVNGVAGPQ-QLPLNGVFLFVSWSRFYQLKPVGTASTARS 540  
Db 616 NGSFDFVIGDSDGSFV---GVSSIPKLEFPFLSASYMI-----QLAKIRLASNIRST 665  
QY 541 L 541  
Db 666 M 666

RESULT 3  
US-09-617-594A-4  
; Sequence 4, Application US/09617594A  
; Patent No. 6541458  
; GENERAL INFORMATION:  
; APPLICANT: Audomnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V  
; FILE REFERENCE: 454313-3151.1  
; CURRENT APPLICATION NUMBER: US/09/617,594A  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 668  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-617-594A-4

Query Match 9.4%; Score 272; DB 4; Length 668;  
Best Local Similarity 28.7%; Pred. No. 6e-18;  
Matches 99; Conservative 39; Mismatches 135; Indels 72; Gaps 15;

QY 21 LVPEVNATA-DP-LPMPFVAGPTTAVATAGQVNMIDPWINVFNVSQGEFTI----- 70  
Db 106 LIGEVAKAMPNLPFLRELEADGDSVTTPQGT-----LVGGVIAEPAQMAVADVATCK 160  
QY 71 -----SPNNTPGDILFDLQPHNLPFLSHLSQMYNGWGNMVRILL 112  
Db 161 SVDSWEAFPSFHTSVNNTSETQKILFKQSLGPLNLPFLSHLSKLYVAVSGSIEVRFS 220



QY 113 LAGNAPSAGKIIVCCVPPGFTS--SSLTIAQATLPHVIAVDTLPIEMPLEDVRNLVH 171  
Db 221 ISGSGVFGKLAIAIVVPPGIDPVQSTMLQ---YPHVLFDAQVPEVIFTIPDLRSLVH 277  
QY 172 TNDNQPTMLVCMWLTPL--RTGGSGNSDSFVVAGRVLTAPSSDFSLFVPTIEQKTR 230  
Db 278 LMSDDTTSLVIMVNDLINPYANDNSNGCIVT--VETKPGDPFKHLLKPP-----GS 330  
QY 231 AFTVNPNIQTLNS-----RFPSLIQMILSPDASQVQFQNGRCLIDQLLGTTPA 283  
Db 331 MLTHGSIPDLIPKSSSLWGNRHWSDIIDFKIP-----FVFAQNRHDFNQ---ETAG 382  
QY 284 TSGQLFRVRGKINQAGARTLNLTVDGKPF---MAFDSAPVGPDP 325  
Db 383 WSTPRFR-----PITTVSEKGGKLGIGVATDVIIP-GIPD 418

RESULT 4  
US-09-590-020-7  
; Sequence 7, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 626  
; TYPE: PRF  
; ORGANISM: Feline calicivirus  
US-09-590-020-7

Query Match 9.3%; Score 268.5; DB 4; Length 626;  
Best Local Similarity 27.5%; Pred. No. 1.2e-17;  
Matches 112; Conservative 45; Mismatches 153; Indels 97; Gaps 18;

QY 10 QSADGASGAGLVPEVNT--ADPLPMEPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEF 68  
Db 79 EADDGSIITA-----PEQGTWGVGVIAEPSSAQMTADMATKGSVDSEW-----EAPF 125  
QY 69 TI-----SPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWGNMRVIRILLAGNAPSAGK 122  
Db 126 SPHTSVNNWSTSETQKILFKQSLGPLNLPYLAKLYVAMSGSIEVRFSISGSGVFGK 185  
QY 123 IIVCCVPPGFTS--SSLTIAQATLPHVIAVDTLPIEMPLEDVRNLVHTNDNQPTML 181  
Db 186 LAAIIVVPPGVDPVQSTMLQ---YPHVLFDAQVPEVIFCLPDLRSLVHLSDDTTSL 242  
QY 182 VCMWLTPL--RTGGSGNSDSFVVAGRVLTAPSSDFSLFVPTIEQKTRAEFTVNIPLQ 240  
Db 243 VIMVNDLINPYANDNSNGCIVT--VETKPGDPFKHLLKPP-----GSMTHGSVPSD 295  
QY 241 TLNS-----RFPSLIQMILSPDASQVQFQNGRCLIDQLLGTTPATSGQLFRVRG 293  
Db 296 LIPKSSSLWGNRHWSDIIDFKIP-----FVFAQNRHDFNQ---ETAGWSTPRFR--- 344  
QY 294 KINQAGARTLNLTVDGKPF---MAFDSAPVGPDPFGKCDWHRISKTPNTSSGDPWRS 350  
Db 345 -----PISVTITEQNGAKLGIGVATDVIIP-GIPD-----GW-----PDTTIPGELI-- 385  
QY 351 VSVQINVQGFVHLGSIQFDEVFNHPTGDYIGTIEWISQPSPTPGTD 397  
Db 386 -----PAGDYAITNGTNDITTATGYD 407

RESULT 5  
US-09-590-020-2  
; Sequence 2, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 623  
; TYPE: PRF  
; ORGANISM: Feline calicivirus  
US-09-590-020-2

Query Match 9.2%; Score 267; DB 4; Length 623;  
Best Local Similarity 31.3%; Pred. No. 1.7e-17;  
Matches 73; Conservative 32; Mismatches 88; Indels 40; Gaps 8;

QY 21 LVPEVNTA-DP-LPMEPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----70  
Db 61 LIARVAEAWDPLDLPLRLEGDDGSITTEQGTM-----VGVVIAEPSSAQMSAADMATOK 115  
QY 71 -----SPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWGNMRVIRIL 112  
Db 116 SVDSWEAEAFSPHTSVNNWSTSETQKILFKQSLGPLNLPYLAKLYVAMSGSIEVRFS 175  
QY 113 LAGNAPSAGKIIVCCVPPGFTS--SSLTIAQATLPHVIAVDTLPIEMPLEDVRNLVH 171  
Db 176 ISGSGVFGKLAIAIVVPPGIEPVQSTMLQ---YPHVLFDAQVPEVIFAIPLRSLVH 232  
QY 172 TNDNQPTMLVCMWLTPL--RTGGSGNSDSFVVAGRVLTAPSSDFSLFVLPVP 223  
Db 233 LMSDDTTSLVIMVNDLINPYANDNSNGCIVT--VETKPGDPFKHLLKPP 283

RESULT 6  
US-09-590-020-4  
; Sequence 4, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 623  
; TYPE: PRF  
; ORGANISM: Feline calicivirus  
US-09-590-020-4

Query Match 9.2%; Score 267; DB 4; Length 623;  
Best Local Similarity 31.3%; Pred. No. 1.7e-17;  
Matches 73; Conservative 32; Mismatches 88; Indels 40; Gaps 8;

QY 21 LVPEVNTA-DP-LPMEPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----70

Db 61 LIAKVAEAWDPLPLFRLEGDDGSIITPEQGT-----VGGVIAEPPSAQMSAAADMATGK 115  
Qy 71 -----SPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWVGNMVRIL 112  
Db 116 SVDSWEAFPSFHTSVNWSSTQKILFKQSLGPLLPYLSHLAKLYVANSQSVVEVRS 175  
Qy 113 LAGNAFSAKIIIVCCVPGFTS-SSLTIAQATLFPFHVIADVRLTEPIEMPLEDVNRVLVH 171  
Db 176 ISGSGVGGKLAIVVPPGIEPVQSTMLQ---YPHVLFDAQVPEVIFAIPDLRSNLVH 232  
Qy 172 TNDNQPTMRLVCMLYTLP-RTGGSGNSDSFVAVAGVLTAPSSDFSLFLVPP 223  
Db 233 LMSDTDTTSLVMVNDLINPYANDTNSGCIVT--VETKPGDPKFKHLKPP 283

RESULT 7  
US-09-590-020-6  
; Sequence 6, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Aivars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 622  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-6

Query Match 8.8%; Score 256; DB 4; Length 622;  
Best Local Similarity 30.9%; Pred. No. 2.1e-16;  
Matches 72; Conservative 33; Mismatches 89; Indels 40; Gaps 8;  
Qy 21 LVEVNTA-DP-LPMEVAGPTTAVATAGQVNMIDPWVNNFVQSGEFTI-----70  
Db 61 LIAKVAEAWDPLPLFRLEGDDGSIITPEQGT-----VGGVIAEPPSAQMSAAADMATGK 115  
Qy 71 -----SPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWVGNMVRIL 112  
Db 116 SVDSWEAFPSFHTSVNWSSTQKILFKQSLGPLLPYLSHLAKLYVANSQSVVEVRS 175  
Qy 113 LAGNAFSAKIIIVCCVPGFTS-SSLTIAQATLFPFHVIADVRLTEPIEMPLEDVNRVLVH 171  
Db 176 ISGSGVGGKLAIVVPPGIEPVQSTMLQ---YPHVLFDAQVPEVIFAIPDLRSNLVH 232  
Qy 172 TNDNQPTMRLVCMLYTLP-RTGGSGNSDSFVAVAGVLTAPSSDFSLFLVPP 223  
Db 233 LMSDTDTTSLVMVNDLINPYANDTNSGCIVT--VETKPGDPKFKHLKPP 283

RESULT 8  
US-07-852-260-2  
; Sequence 2, Application US/07852260  
; Patent No. 5525715  
; GENERAL INFORMATION:  
; APPLICANT: Racaniello, Vincent  
; APPLICANT: Tate, Joanne M.  
; APPLICANT: Weeks-Levy, Carolyn L.  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM  
; TITLE OF INVENTION: CDNA  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham

STREET: 30 Rockefeller Plaza  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10112  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/852,260  
FILING DATE: 19920619  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 36607-B-PCT-US  
TELEPHONE: (212) 977-9550  
TELEFAX: (212) 664-0525  
TELEX: 422523 COOP UI  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2206 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-07-852-260-2  
Query Match 4.8%; Score 140; DB 1; Length 2206;  
Best Local Similarity 21.2%; Pred. No. 0.0006;  
Matches 122; Conservative 62; Mismatches 221; Indels 170; Gaps 27;  
Qy 17 GAGQLVPEVNTADPLPMEPVAGPTTAVATAGV-NM-----IDPWVNNFVQSPQG---66  
Db 349 GSNQYLTSDNHQSPCAI-PEFDVTPPIDIPGEVKNMMLAEIDTMIPLNLESTKENTMDM 407  
Qy 67 -EFTISPNNTPGDILFDLQGLPHLPFLSH-----LSQMYNGWVGNMVRILLAGNAFSA 120  
Db 408 YRVLTSDSADLSQPLCLSLSPAFDPRLSHTMLGVLNYYTHWAGSLKFTFLFCGSMAT 467  
Qy 121 GKIIIVCCVPPGFTSSLTIAQATLFPFHVIADVRLTEPIEMPLEDVNRVLVH--TNDNQPT 178  
Db 468 GKILVAYAPPG-ADPPTSRKEAMLGTHVINDLGLQSSCTMVVPMWISNTYRQTQDSFTE 526  
Qy 179 MRLVCMLY-----TPLRTGGSGNSDSFVAVAGVLTAPSSDFSLFLVPPTIEQKRAFT 233  
Db 527 GGYISNFTQTRIVVPLST-----PKSMSMLGFV--SACNDFSVRLLRDTT-----569  
Qy 234 VPMIPLQTLNSRFPPLSIQGMILSPDASQVVQFQNGRCLIDGQLLGTTPATSGQLFRVRG 293  
Db 570 -----HISQSALPQIEDL-----TS 585  
Qy 294 KINQARTLNL-TEVDGKPFNAFDSAPVGPDPFGKCDHMR--ISKTPNNTSSGDPMR- 349  
Db 586 EVAQGALTLSLPKQDLSLFDTKASGPA-----HSKEVPALTAVETGATNPLAP 633  
Qy 350 SVSVQT-----NVQGFVPHLGSIQDFEVNHPHTGDIYIGTIEWISQSPSTPPGTD 397  
Db 634 SDTVQTRHVQRSSRSESTIESFFARGACVAILVDN-----EQPTTRAKQL 680  
Qy 398 INLWEIPDYSSLSQAANLAPPVFPFGGEALVYFSA-FPGPNRNSAPNDVPCLLPQBY 456  
Db 681 FAWMRI-TYKDTVQLRRKLEFFTY-SRFDMEFTFVVNTANNGHALNQV-----Y 731  
Qy 457 ITHFVSEQAPTMGDAALLHVDP-----DTNRLNGEFLKYPGGYLTCPVNGVGCAGPQ 508  
Db 732 QIMYIPPGAPT-----PKSWDDYTWTQTSNPSIFTY-----GAAFA 768  
Qy 509 QLPNGVFLFVSWVSRFYQ-LKPVGTGASTARSRLG 542  
Db 769 RISVPVGLANAY-SHFYDGFAGVPLKTDANDQIG 802

RESULT 9  
US-08-461-503-2  
; Sequence 2, Application US/08461503  
; Patent No. 5834302  
; GENERAL INFORMATION:  
; APPLICANT: Racaniello, Vincent  
; APPLICANT: Tatem, Joanne M.  
; APPLICANT: Weeks-Levy, Carolyn L.  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES  
; TITLE OF INVENTION: FROM CDNA  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,503  
; FILING DATE: 5-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2206 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-461-503-2

Query Match 4.8%; Score 140; DB 2; Length 2206;  
Best Local Similarity 21.2%; Pred. No. 0.0006;  
Matches 122; Conservative 62; Mismatches 221; Indels 170; Gaps 27;  
QY 17 GAGQLVPEVNTADPLPMEPVAGTTAVATAGV-NM-----IDPMVNNFVQSPQ----- 66  
DB 349 GSNQYLTSNDHQSPCAI-PEFDVTPPIDPGEVKMMELAEIDTWPLNLESTKRTNMDM 407  
QY 67 -EFTISPNNTPGDILDLQGLPHLNPFLSH-----LSQMYNGWGNMVRILLAGNAPSA 120  
DB 408 YRVTLSDSADLSQPIILCLSLSPAFDRLSHHTMLGEVLYNTHWAGSLKFTFLFCGSMMAT 467  
QY 121 GKIIIVCCVPPGTSSTLTIAQATLPHVIADVRLTLEPIEMPLEDVNRVLYH--TNDNQPT 178  
DB 468 GKILVAYAPPG-AQPTSKKAWLGTHTVINDLGLOSSCTMNVVFNISNVTYRTQDSFTE 526  
QY 179 MRLVCMVLY-----TPLRTGGSGNSDSFVAVGRVLTAPSSDFSFLFVLPPTIEQKTRAPT 233  
DB 527 GGYISMFYQTRIVVPLST-----PKSMMLGFV--SACNDFSRLLRDTT----- 569  
QY 234 VNIPIQLTISNRFPISLIGMILSPDASQVQFQNGRCLIDGQLLGTTPATSQLFRVNG 293  
DB 570 -----HISQALPQGIEDL-----TS 585  
QY 294 KINQARTLNL-TEVDGKPFMAFDSPAPVFPDFGKCDWHR--ISKTPNNTSSGDPMR- 349  
DB 586 EVAQGALTLSLPKQDLSLPTKASGPA-----HSKEVPALTAVETGATNPLAP 633

QY 350 SVSVQVT-----NVQGFVPHLGSIQPDEVFNHPTGDIYCTIEWISQSPSTPPGTD 397  
DB 634 SDTVQTRHVQRRSRSESTIESFFARGACVAIIEVDN-----EQFTTRAQKL 680  
QY 398 INLWEIPDYGSSLSQAANLAPVPFPFGGEALVYFVSA-FPGPNNSAPNDVPCLLPQBY 456  
DB 681 FAWMRI-TYKTVQURRKLEFFTY-SRDMETTFVTANTFTNANGHALNQV-----Y 731  
QY 457 ITHFVSEQAPTMGDAALLHYVDP-----DTNRNLGEFKLYPGGYLTCVPNGVGAGPQ 508  
DB 732 QIMYIPPGAPT-----PKSWDDVYTWQTSNPSIFTY-----GAAPA 768  
QY 509 QLPNGVFLFVSWWSRFYQ-LKPVGCTASTASRLG 542  
DB 769 RISVPYVGLANAY-SHFYDGFAPKVLKTDANDQIG 802  
RESULT 10  
US-08-465-250-2  
; Sequence 2, Application US/08465250  
; Patent No. 6136570  
; GENERAL INFORMATION:  
; APPLICANT: Racaniello, Vincent  
; APPLICANT: Tatem, Joanne M.  
; APPLICANT: Weeks-Levy, Carolyn L.  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM  
; TITLE OF INVENTION: CDNA  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release 1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,250  
; FILING DATE: 6-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2206 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-465-250-2  
Query Match 4.8%; Score 140; DB 3; Length 2206;  
Best Local Similarity 21.2%; Pred. No. 0.0006;  
Matches 122; Conservative 62; Mismatches 221; Indels 170; Gaps 27;  
QY 17 GAGQLVPEVNTADPLPMEPVAGTTAVATAGV-NM-----IDPMVNNFVQSPQ----- 66  
DB 349 GSNQYLTSNDHQSPCAI-PEFDVTPPIDPGEVKMMELAEIDTWPLNLESTKRTNMDM 407  
QY 67 -EFTISPNNTPGDILDLQGLPHLNPFLSH-----LSQMYNGWGNMVRILLAGNAPSA 120  
DB 408 YRVTLSDSADLSQPIILCLSLSPAFDRLSHHTMLGEVLYNTHWAGSLKFTFLFCGSMMAT 467  
QY 121 GKIIIVCCVPPGTSSTLTIAQATLPHVIADVRLTLEPIEMPLEDVNRVLYH--TNDNQPT 178

Db 468 GKILVAYAPG-AQPPTSKEAMLGTHVWDLGLQSSCTMVVPMISNVYTRQTQDSFTE 526  
QY 179 MELVCMYX-----TFLRTGGSGNSDSFVAVAGRVLTAPSSDFSFLFLVPPTIEQKTRAF 233  
Db 527 GGYISMFYQTRIVVPLST-----PKSMMLGFV--SACNDFSVALLRDIT-- 569  
QY 234 VNIPIQLTNSRFPSLIQGMILSPDASQVQFQNGRCLIDQLLGTTPATSGQLFRVRG 293  
Db 570 -----HISQALFQGIEDL----- 585  
QY 294 KINQOGARTLNL-TEVDGKPFMAFDSPAPVPFDFGKCDWHMR--ISKTPNNTSSGDPNR- 349  
Db 586 EVAQGALTSLPKQODSLPDKASGPA-----HSKEVPALTAVETGATNPLAP 633  
QY 350 SVSVQV-----NVQGFVHLGSIQDFEVENHPTGDIYGTIEWISOPSPPTGTD 397  
Db 634 SDTVQTRHVQRRSRSESTIESFFARGACVALLIEDN-----EQPTTRAQKL 680  
QY 398 INLWEIPDYGSSLSQAAANLAPPVFPFGGEALVYFVSA-FPGNNRSAPNDVPCLLPQBY 456  
Db 681 FAWMRI-TYKDTVQLRRKLEFFTY-SRFDMEFTFVVTANFTNANGHALNQV-----Y 731  
QY 457 ITHFVSEQAPTMGDALLHHVDP-----DTNRNLGEFKLYPGGYLTCVENVGVGAGPQ 508  
Db 732 QIMYIPPGAPT-----PKSWDDYTWTSSNPSIFYTY-----GAAPA 768  
QY 509 QLPGLNGVELFVMSVSRFYQ-LKPVGTASTARSRLG 542  
Db 769 RISVPVGLANAY-SHYDGFAGVPLKTDANDQIG 802

RESULT 11  
US-09-579-181-2  
; Sequence 2, Application US/09579181  
; Patent No. 6365372  
; GENERAL INFORMATION:  
; APPLICANT: Chrivia, John  
; APPLICANT: Yaciuk, Peter  
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)  
; FILE REFERENCE: 16153-4247  
; CURRENT APPLICATION NUMBER: US/09/579,181  
; CURRENT FILING DATE: 2000-05-25  
; PRIOR FILING DATE: 1999-05-27  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 2972  
; TYPE: PRT  
; ORGANISM: Human  
US-09-579-181-2

Query Match 4.2%; Score 123; DB 4; Length 2972;  
Best Local Similarity 21.1%; Pred. No. 0.048;  
Matches 122; Conservative 57; Mismatches 206; Indels 192; Gaps 26;  
QY 19 GOLVPEVNT----ADPLEMEVAG-TTAVATAGQVNMIDP--WIVNMFVQSPQGETTIS 71  
Db 809 GRTVVVNNPRAPLGPVDPVPPGPELSAQTPGVPQVLPASLMVS---ASPAGPPLIP 865  
QY 72 PNNTFGDILFDLQGLPHLNPFLSHLSOMYNGWGNMR-----VRIL 112  
Db 866 ASRPPGPVLL-----PPLQPNSGSLFQVLPGLVSLGTSRPPPTLSLKEPTTAPVRLS 920  
QY 113 LAGNAFSAAGKIIVCCVPPGF-----TSSSLTTIAQATLFPFHVIADVRLTPIEMPLEDV 165  
Db 921 PAPPGSSSLKPLTVPGYTFPPAAATTTTATAT-----TTAVPAPTAPQ- 970  
QY 166 RNVLVHTNDNQPM-----RLVCMLYTPLTRGGSGNSDSFVAVAGRVLTAPSSDFS 216  
Db 971 RLIL--SPDMQARLPSGEVVSIGQLASLAQRVANAGGS-KPLTFQIQGNKLTLTGAQVR 1027

QY 217 FLFL-----VPPTIEQKTRAFV-----PNIPLOTLSNSRF 247  
Db 1028 QLAVGQPRFLQHPPTWNTGVKIVVROAPRDGLTPVPPLAPAPRPSSGGLPAVLNPR- 1086  
QY 248 PSLIQGMILSPDASQVQFQNGRCLIDQLLGT--TPATSGQLFRVRGKINGARTLNL 305  
Db 1087 PTLTPGRLPTP-----TLGTARAPMPTPTLVRPLLLKLVHSPSPPEVSA 1128  
QY 306 EVDGKPFMAFDSP--APVGFDFGKCDWHMRISKTPNNTSSGDPNRSVSQVQTNVQGFVPH 363  
Db 1129 SAFGAAPLTITSSPLHVPSSLP--GPASSMPPI--PNSSPLASPVSS-TVSPLSSSLP- 1181  
QY 364 LGSIQDFEVENHPTGDIYGTIEWISOPSTPPGTDINLWEIPDYGSSLSQAA-----NL 416  
Db 1182 -----ISVPTTLPAASAPLTIP-ISAPLTVSASGPALLTSV 1217  
QY 417 APPVFP-----PGFGEALVYFVSAFPG 438  
Db 1218 TPPLAPVVPAAAPGPSLQPSGASASALTGLATAPSLSSSQTPGHPDLLLAPTSSHVPG 1277  
QY 439 PNNRSAPNDVPCLLPQBYI-THFVSEQAPTMGDAALL 474  
Db 1278 LNSTVAPACSPVLVPASALASPPSPAPNAPAPQASILL 1314

RESULT 12  
US-09-579-181-1  
; Sequence 1, Application US/09579181  
; Patent No. 6365372  
; GENERAL INFORMATION:  
; APPLICANT: Chrivia, John  
; APPLICANT: Yaciuk, Peter  
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)  
; FILE REFERENCE: 16153-4247  
; CURRENT APPLICATION NUMBER: US/09/579,181  
; CURRENT FILING DATE: 2000-05-25  
; PRIOR FILING DATE: 1999-05-27  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3118  
; TYPE: PRT  
; ORGANISM: Human  
US-09-579-181-1

Query Match 4.2%; Score 123; DB 4; Length 3118;  
Best Local Similarity 21.1%; Pred. No. 0.052;  
Matches 122; Conservative 57; Mismatches 206; Indels 192; Gaps 26;  
QY 19 GOLVPEVNT----ADPLEMEVAG-TTAVATAGQVNMIDP--WIVNMFVQSPQGETTIS 71  
Db 955 GRTVVVNNPRAPLGPVDPVPPGPELSAQTPGVPQVLPASLMVS---ASPAGPPLIP 1011  
QY 72 PNNTFGDILFDLQGLPHLNPFLSHLSOMYNGWGNMR-----VRIL 112  
Db 1012 ASRPPGPVLL-----PPLQPNSGSLFQVLPGLVSLGTSRPPPTLSLKEPTTAPVRLS 1066  
QY 113 LAGNAFSAAGKIIVCCVPPGF-----TSSSLTTIAQATLFPFHVIADVRLTPIEMPLEDV 165  
Db 1067 PAPPGSSSLKPLTVPGYTFPPAAATTTTATAT-----TTAVPAPTAPQ- 1116  
QY 166 RNVLVHTNDNQPM-----RLVCMLYTPLTRGGSGNSDSFVAVAGRVLTAPSSDFS 216  
Db 1117 RLIL--SPDMQARLPSGEVVSIGQLASLAQRVANAGGS-KPLTFQIQGNKLTLTGAQVR 1173  
QY 217 FLFL-----VPPTIEQKTRAFV-----PNIPLOTLSNSRF 247  
Db 1174 QLAVGQPRFLQHPPTWNTGVKIVVROAPRDGLTPVPPLAPAPRPSSGGLPAVLNPR- 1232  
QY 248 PSLIQGMILSPDASQVQFQNGRCLIDQLLGT--TPATSGQLFRVRGKINGARTLNL 305  
Db 1233 PTLTPGRLPTP-----TLGTARAPMPTPTLVRPLLLKLVHSPSPPEVSA 1274



Wed Jun 2 09:13:27 2004

us-09-926-799-3.ra1

Search completed: June 1, 2004, 13:57:56  
Job time : 15.0087 secs

```
Db      2599 NVGTGDHV 2606
| |||::
| |||::

RESULT 15
US-09-976-594-15
; Sequence 15, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS
; FILE REFERENCE: PA-0041 US
; CURRENT APPLICATION NUMBER: US/09/976,594
; CURRENT FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: 60/240,409
; PRIOR FILING DATE: 2000-10-12
; NUMBER OF SEQ ID NOS: 1143
; SOFTWARE: PERL Program
; SEQ ID NO 15
; LENGTH: 2322
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. 6673549 1720920CD1
US-09-976-594-15

Query Match      3.88; Score 109; DB 4; Length 2322;
Best Local Similarity 22.08; Pred. No. 0.79;
Matches 131; Conservative 69; Mismatches 193; Indels 202; Gaps 35;

Qy      10 QSADGASGAG-----LVPEVNTADPLPMEFVAGPTT-----AVATAGQVNMIDPWIV 57
Db      1805 QGPAGASVAGPQTSFAITVRDNERPPQASVPLRLTRGSRAPISRAQLSVVD-----1860
Qy      58 NNVSQSGEFTTSPNTPGDIILFDLQGLHNPFLSHLSQMYNGWGNM-----R 108
Db      1861 -----PDSAPGEIEYEVORAPH-NGFLS-----LVGGGLGPVTRFTQADVDS 1901
Qy      109 VRILLAGNAGSAGIIVCCVPPGPTSSSLTIAQAT-LFPHVIADVRTLEPIEMPLEDVNRN 167
Db      1902 GRFAVANGSVAGIFQLSMDG-ASPLPMSLAVDILPSAI-EVQLRAPLEVP-----Q 1954
Qy      168 VLYHTNDNQPMRLVCMLYTP-----LRTGGGNSDSFVVGREVLTAPESSDFLFLVP 222
Db      1955 ALGRSSLSQQQLRVVSDREEPEAAAYRLIQGPYGH---LLVGGR-----PTSASFQF---2003
Qy      223 PTIEQKTAFTVPIPILOTLSNRFPSSLIOGMILSPDASQVQF-----Q 267
Db      2004 -QIDQGEVVFAFTNFS-SSHDFRVLALRGV-----NASAVNVTVRALLHWAGGPWQ 2057
Qy      268 NGRCLIDGQLL-----GTTP-----ATGQLFRV-RGKINQAGTL--NLTEVD 308
Db      2058 GATLRLDPTVLDAELANRTGSPFRLLGPRHGRVVRVPRARTEFGGSQLVEQFTQD 2117
Qy      309 -----GKPFMAFSPAPVGPDPFGKCDMHRISKTPNTSSGDPMRSVSVQTNVQ 358
Db      2118 LEDGRGLGLEVRP--EGRAPGAG-----DSLLELWQ 2149
Qy      359 GFVPHLGSIQFDEVFNHPTGDYIGTIEWISOP-----STPGTDINLWEIPD- 405
Db      2150 GUPPAVASLDFAPEYNAARPY--SVALLSVPEAAATEAGKPESSTPTGEPGPMASSEP 2207
Qy      406 ---YGSLSLQ-QAANLAPPVFPFGFGEALVYFVSAPFGP-----NNESAPNDVPCLLPQ 454
Db      2208 AVAKGGFLSFLKANMFSVIIP-----MCLVLLALLILPLLFLYLRKRNTKGRHDVQVLTAK 2263
Qy      455 EVITHFVSQAPTMGDAALLHVVDPTNRLNGLKFLYPCGY--LTCVFNQVGVGAP 507
Db      2264 -----PRNGLAG-----DTETFR-----KVEFGQAIPLTAVP---GQGP 2294
```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.9037 Seconds  
(without alignments)  
4368.312 Million cell updates/sec

Title: US-09-926-799-3

Perfect score: 2895

Sequence: 1 MMASKDAPQSDAGAGAGQ.....QLKPVGTASTARSLGVRR 546

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2034.5	70.3	530	14	US-10-314-739-3
2	290	10.0	669	14	US-10-209-507-2
3	272	9.4	668	14	US-10-209-507-4
4	268.5	9.3	671	12	US-10-670-695-2
5	268	9.3	547	12	US-10-670-695-4
6	132.5	4.6	1194	12	US-10-282-122A-46163
7	127.5	4.4	2126	15	US-10-052-648A-39
8	123.5	4.3	1168	15	US-10-369-493-16733
9	123	4.2	2971	14	US-10-146-473-50
10	122.5	4.2	6310	12	US-10-282-122A-67793
11	115.5	4.0	1765	12	US-10-282-122A-48055
12	114.5	4.0	1194	12	US-10-282-122A-46577
13	113.5	3.9	1147	15	US-10-327-481A-38
14	111	3.8	1046	14	US-10-224-999A-3480
15	110.5	3.8	1394	12	US-10-381-247B-18

16	110.5	3.8	1435	12	US-10-276-774-2178
17	110	3.8	896	14	US-10-205-342-27
18	109.5	3.8	2468	12	US-10-282-122A-66335
19	109.5	3.8	2468	14	US-10-246-330-4
20	109	3.8	2322	10	US-09-919-039-15
21	108.5	3.7	1025	11	US-09-834-309-5
22	108.5	3.7	1106	14	US-10-157-031-30
23	108	3.7	816	12	US-10-425-114-55912
24	107	3.7	1048	14	US-10-223-538-10
25	106.5	3.7	3399	15	US-10-080-334-196
26	106	3.7	1426	12	US-10-664-859-15
27	106	3.7	1426	12	US-09-915-543-15
28	106	3.7	1426	14	US-10-322-579-15
29	105	3.6	575	9	US-09-738-626-4263
30	104.5	3.6	2514	12	US-10-282-122A-66121
31	104.5	3.6	2514	15	US-10-320-800-40
32	102.5	3.5	2135	12	US-10-362-892-9
33	102.5	3.5	2135	15	US-10-288-798-9
34	102.5	3.5	2382	12	US-10-336-472-230
35	102.5	3.5	2382	14	US-10-136-935A-2
36	102.5	3.5	2382	15	US-10-052-648A-40
37	102.5	3.5	3354	15	US-10-080-334-197
38	102	3.5	944	14	US-10-174-677-101
39	102	3.5	1185	15	US-10-259-194A-246
40	102	3.5	2358	12	US-10-282-122A-45763
41	102	3.5	3930	12	US-10-282-122A-46817
42	101.5	3.5	680	16	US-10-389-566-429
43	101.5	3.5	752	12	US-10-425-114-71415
44	101.5	3.5	1491	12	US-10-282-122A-65570
45	101.5	3.5	3013	12	US-10-282-122A-78257

#### ALIGNMENTS

#### RESULT 1

US-10-314-739-3  
Sequence 3, Appli US/10314739  
Publication No. US20030129588A1  
GENERAL INFORMATION:  
APPLICANT: Estes, Mary K  
Jiang, Xi  
Graham, David Y  
TITLE OF INVENTION: Methods and Reagents to Detect and Characterize No. US20030129588A1walk and Related Viruses  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski L.L.P.  
CITY: Washington, D.C.  
STATE: <Unknown>  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA: US/10/314,739  
APPLICATION NUMBER: US/10/314,739  
FILING DATE: 09-Dec-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,049  
FILING DATE: June 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Davis, Peter  
REGISTRATION NUMBER: 36,119  
REFERENCE/DOCKET NUMBER: 311.023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-662-0200  
TELEFAX: 202-662-4643  
TELEX: <Unknown>

Sequence 2178, Appli  
Sequence 27, Appli  
Sequence 66335, A  
Sequence 4, Appli  
Sequence 15, Appli  
Sequence 5, Appli  
Sequence 30, Appli  
Sequence 55912, A  
Sequence 10, Appli  
Sequence 196, Appli  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 15, Appli  
Sequence 4263, Ap  
Sequence 66121, A  
Sequence 9, Appli  
Sequence 9, Appli  
Sequence 2, Appli  
Sequence 230, Appli  
Sequence 197, Appli  
Sequence 101, Appli  
Sequence 246, Appli  
Sequence 45763, A  
Sequence 46817, A  
Sequence 429, Appli  
Sequence 71415, A  
Sequence 65570, A  
Sequence 78257, A

; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-314-739-3

Query Match 70.3%; Score 2034.5; DB 14; Length 530;  
Best Local Similarity 69.4%; Pred. No. 7.2e-183; Indels 17; Gaps 6;  
Matches 379; Conservative 69; Mismatches 81;

QY 1 MMWASKDAPOSAGAGQQLVPEVNTADPLMEPVAGPTTAVATAGQVNMIDPIMVNF 60  
DB 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPIMNF 60  
QY 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLNPFSLHLSOMYNGWGNMVRILLAGNAFSA 120  
DB 61 VQAPQGEFTISPNNTPGDVLFDLSGLPHLNPFLLHLSOMYNGWGNMVRIMLAGNAFTA 120  
QY 121 GKIIICVCPGFTSSSLITIAQATLFPFHVIAVRTLEPIEMPLEDVNRNLYHTND--NQPTM 179  
DB 121 GKIIICVCPGFTSSSLITIAQATLFPFHVIAVRTLEPIEMPLEDVNRNLYHTND--NQPTM 180  
QY 180 RLVCMLYTPLRITGGSGNSDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVNPIL 239  
DB 181 RLVCMLYTPLRITGGGTG--DSFVAGRVMTCPSPDENFLFVLPPTVEQKTRFPLNPL 238  
QY 240 QTLNSRFPPLSQMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGA 299  
DB 239 SSSLNSRAPLPISSMGISPDNVQSVQFQNGRCLIDGRLVGTTPVSLHAKINGTSN--G 296  
QY 300 RTNLNTEVDCGKPFMAPDSPAPVGPDPFGKCDMHRISKTPNNTSGDPMRSVSQVNTVQ 359  
DB 297 TVINLTEDGTTPHFPFGCPAPIGPDGLGGCDMHNMTQFGHSSQT-----QYDVTTPDT 351  
QY 360 FVPHLSIQFDEVFNHPTGDIYGTIEWISQSPTPPGTDINLWEIPDYGSSLSAANLAPP 419  
DB 352 FVPHLSIQANGI---GSGNVGVLSWISPPHSGSQVDLWKIPNYSITTEATHLAFS 408  
QY 420 VFPFGCEALYFVFSAPGPNRSAPNDVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDP 479  
DB 409 VYPPFGGEVLVFFMSKMPG-----GAYNLPCLLPQEIYTHFVSEQAPTMGDAALLHYVDP 464  
QY 480 DTNRNLGEFKLYPGGYLTCVPNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARS 539  
DB 465 DTGRNLGEFKAYPDGFLTCVPNGASSGPGQQLPINGVFVFSWVSRYQLKPVGTASSARG 524  
QY 540 RLGVRR 545  
DB 525 RLGLRR 530

RESULT 2  
US-10-209-507-2  
; Sequence 2, Application US/10209507  
; Publication No. US20030109033A1  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.2  
; CURRENT APPLICATION NUMBER: US/10/209,507  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: 09/617,594  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 669  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-10-209-507-2

Query Match 10.0%; Score 290; DB 14; Length 669;  
Best Local Similarity 25.5%; Pred. No. 5.3e-18;  
Matches 153; Conservative 71; Mismatches 247; Indels 130; Gaps 29;

QY 12 ADGAGAGQQLVPEVNT--ADPLMEPVAGPTTAVATAGQVNMIDPIMVNFVQSPQGEFTI 70  
DB 125 ADGDSI--TTPGQTLVGVTAEPSAOMATAATAATGKSDVSEW--ESFF-SFHTSVNM 179  
QY 71 SPNTPGDIILFDLQGLPHLNPFSLHLSOMYNGWGNMVRILLAGNAFSAAGKIIIVCCVPP 130  
DB 180 STSETQKILFKOSLGLPFLNPFLEHLSKLYVAVSGSDVRFISGSGVGGKLAIVVPP 239  
QY 131 GFTS--SSLITIAQATLFPFHVIAVRTLEPIEMPLEDVNRNLYHTNDNQPTMRLVCMLYTPL 189  
DB 240 GVDPVQSTMLQ--YPHVLFDARQVEPVIFSPDLRSTLYHLSMDTDTTSLVIMYNDL 296  
QY 190 -RTGGSGNSDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVNPILQTLNS-- 245  
DB 297 INPYANDSNSGGCIVT--VETKPGDPFKFLLKPP-----GSMLTGHSIPSDILPKSSSL 349  
QY 246 ----RPFSLIQMILSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGA 301  
DB 350 WIGNRYWSDITDFVIRP-----FVFQNRHFDNQ---ETAGWSTPRFR-----PIT 393  
QY 302 LNLTEVDGKPF--MAFDSAPVGPDPFGKCDMHRISKTP-----NNTSGDP 347  
DB 394 ITISESGSKLGTGATVDIYV-GIPD-GWPDITIGBELTPAGDYITNGSGNDIATANA 451  
QY 348 MRSVSQVNTVQGVFPHLSIQFDEVFNHPTGDIY--GTIE--W-----ISQP 390  
DB 452 YDSADVTITNTNF-----RGMVYICGALQRAWGDKKISSATFITTALKEG 495  
QY 391 ST-PEGTDINLWEIPDY-----GSSLSQAANLAPPVFPFGFGEAL-----VYFVSAFP 437  
DB 496 NTLKPSNTIDMTKLVQVQTHVGRDVTSDDTLAILGYVTGIGEQAIGSNRDSVVRISMLP 555  
QY 438 GPNRNSA-----PNDVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDP-----DT 481  
DB 556 ETGARGGNHPIFYKNSIKLGYVLSIDVFNLSQILHTRQJSLNHYLLPPDSFAVYRIDS 615  
QY 482 NRNLGEFKLYPGGYLTCVPNGVAGPQ-QLPLNGVFLFVSWSRFYQLKPVGTASTARS 540  
DB 616 NGSWFDVGIDSDGFSFV---GVSSIPKLEFPLSASYMGI-----QLAKIRLASNIRST 665  
QY 541 L 541  
DB 666 M 666

RESULT 3  
US-10-209-507-4  
; Sequence 4, Application US/10209507  
; Publication No. US20030109033A1  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.2  
; CURRENT APPLICATION NUMBER: US/10/209,507  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: 09/617,594  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421



```
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-4

Query Match          9.4%; Score 272; DB 14; Length 668;
Best Local Similarity 28.7%; Pred. No. 2.7e-16;
Matches 99; Conservative 39; Mismatches 135; Indels 72; Gaps 15;

QY 21 LVEPVNTA-DP-LPMEPVAGPTTAVATAGQVNMIDPWVNNVQSPQGEFTI-----70
Db 106 LIGEVAKADPNLPLFRLEADGGSVTPRQGT-----LVGGVIAEPNAQMSAVADVATGK 160
QY 71 -----SPNPTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMRVRIL 112
Db 161 SVDSWEAPFSSHTSVNNSTSTQKILFKQSLGPLNPLYLHLAKLYVAMSGSIEVRPS 220
QY 113 LAGNAFSAKIIIVCCVPPGFTS-SSLTIAQAATLPHVIAADVRLTLEPIEMPLEDVNRNLYH 171
Db 221 ISGSGVFGKLAIVVPPGIDFPVQSTMLQ----YHVLFDARQVEPVIFITPDLNSLYH 277
QY 172 TNDNQPTMRLVCMLYTPL-RTGGSGNSDSFVAGRVLTAPSSDFSFLVLPPTIEQKTR 230
Db 278 LMSDTTDTSLVIMYNDLINPYANDSNSSGCIVT--VETKPGDPFKFHLKPP-----GS 330
QY 231 AFTVNPILQTLNS-----RPSLIQGMILSPDASQVVOFQNGRCLIDGQLLGTTPA 283
Db 331 MUTHGSIPLSLIPKSSSLWGNRHSIDITDFVIKP-----FVFOANRHDFNQ----ETAG 382
QY 284 TSGQLFRVRGKINQKARTLNLEVDGKPF---MAFDSAPAPVGFDP 325
Db 383 WSTPRP-----PIITVSEKSGKLGIGVATDSIVP-GIPD 418

RESULT 4
US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; PRIOR FILING DATE: 2003-09-25
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match          9.3%; Score 268.5; DB 12; Length 671;
Best Local Similarity 26.8%; Pred. No. 5.7e-16;
Matches 110; Conservative 47; Mismatches 149; Indels 105; Gaps 18;

QY 10 QSADGASGAGQLVPEVNT-ADPLPMEPVAGPTTAVATAGQVNMIDPWVNNVQSPQGEF 68
Db 124 EADDGSITA---PEQGTWVGVIAPSAQMSQMSAADMATGKSVDSSEW-----EAPF 170
QY 69 TI-----SPNPTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMRVRILLAGNAFSAK 122
Db 171 SPHTSVNNSTSTQKILFKQSLGPLNPLYLHLAKLYVAMSGSIEVRPSISGSGVFGK 230

; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-4

Query Match          9.4%; Score 272; DB 14; Length 668;
Best Local Similarity 28.7%; Pred. No. 2.7e-16;
Matches 99; Conservative 39; Mismatches 135; Indels 72; Gaps 15;

QY 21 LVEPVNTA-DP-LPMEPVAGPTTAVATAGQVNMIDPWVNNVQSPQGEFTI-----70
Db 106 LIGEVAKADPNLPLFRLEADGGSVTPRQGT-----LVGGVIAEPNAQMSAVADVATGK 160
QY 71 -----SPNPTPGDILFDLQGLPHLNPFLSHLSOMYNGWGNMRVRIL 112
Db 161 SVDSWEAPFSSHTSVNNSTSTQKILFKQSLGPLNPLYLHLAKLYVAMSGSIEVRPS 220
QY 113 LAGNAFSAKIIIVCCVPPGFTS-SSLTIAQAATLPHVIAADVRLTLEPIEMPLEDVNRNLYH 171
Db 221 ISGSGVFGKLAIVVPPGIDFPVQSTMLQ----YHVLFDARQVEPVIFITPDLNSLYH 277
QY 172 TNDNQPTMRLVCMLYTPL-RTGGSGNSDSFVAGRVLTAPSSDFSFLVLPPTIEQKTR 230
Db 278 LMSDTTDTSLVIMYNDLINPYANDSNSSGCIVT--VETKPGDPFKFHLKPP-----GS 330
QY 231 AFTVNPILQTLNS-----RPSLIQGMILSPDASQVVOFQNGRCLIDGQLLGTTPA 283
Db 331 MUTHGSIPLSLIPKSSSLWGNRHSIDITDFVIKP-----FVFOANRHDFNQ----ETAG 382
QY 284 TSGQLFRVRGKINQKARTLNLEVDGKPF---MAFDSAPAPVGFDP 325
Db 383 WSTPRP-----PIITVSEKSGKLGIGVATDSIVP-GIPD 418

RESULT 4
US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; PRIOR FILING DATE: 2003-09-25
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2

Query Match          9.3%; Score 268; DB 12; Length 547;
Best Local Similarity 26.9%; Pred. No. 4.7e-16;
Matches 104; Conservative 43; Mismatches 139; Indels 100; Gaps 16;

QY 34 BPVAGPTTAVATAGQVNMIDPWVNNVQSPQGEFTI-----SPNPTPGDILFDLQGLP 87
Db 21 EPSAQMSAADMATGKSVDSSEW-----EAFSFTSVNNWSTSETQKILFKQSLP 71
QY 88 HLNPLSHLSOMYNGWGNMRVRILLAGNAFSAKIIIVCCVPPGFTS-SSLTIAQAATLPP 146
Db 72 LLNPYLEHLAKLYVAMSGSIEVRPSISGSGVFGKLAIVVPPGVDVQSTMLQ---YP 128
QY 147 HVIADVRLTLEPIEMPLEDVNRNLYHTNDNQPTMRLVCMLYTPL-RTGGSGNSDSFVAG 205
Db 129 HVLFDARQVEPVIFCLPDLRSTLYHLSMDTDTTSLVIMYNDLINPYANDSNSSGCIVT- 187
QY 206 RVLTAPESSDFSFLVLP-----TTEQKTRAFVNPILQTLNSRFPSLIQM 254
Db 188 -VETKPGDPFKFHLKPPGSMITHGSIPLSLIPKTSLSLWGN-----RYWSDITDF 237
QY 255 ILSPDASQVVOFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQKARTLNLEVDGKPF-- 312
Db 238 VIRP-----FVFOANRHDFNQ----ETAGWSTPRF-----PISVTITEQNGAKLGI 281
QY 313 -MAFDSAPAPVGFDPFKGKCDWHMRISKTNNNTSSGDPMSVSVQTNVQGFVPHLSIQPDE 371
Db 282 GVATDYIYP-GIPD-----GW-----PDTTIPGELI----- 306
```

Db 372 VFNHPTGDIYGTIEWISQSPSTPGTD 397  
Db 307 ----PAGDYAITNGTGNDDITATGYD 328

RESULT 6  
US-10-282-122A-46163  
; Sequence 46163, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 46163  
; LENGTH: 1194  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
; FEATURE:  
; NAME/KEY: MISC\_FEATURE  
; LOCATION: (3)..(3)  
; OTHER INFORMATION: X=any amino acid  
US-10-282-122A-46163

Query Match 4.6%; Score 132.5; DB 12; Length 1194;  
Best Local Similarity 21.0%; Pred. No. 0.0097;  
Matches 126; Conservative 60; Mismatches 203; Indels 211; Gaps 30;

QY 19 GOLVPEVN---TADPLMEFVAGPTTAVATACQVN-MIDPWTVN-NFVQSPQ---GB-- 67  
Db 242 GTITNEANVTYSQNPTEP---PTTTTTPPTNTSVRTALVNPFTKTVSFQVADIGDII 298  
QY 68 -FTISPNNT-----PGDILFDLQGLPHLPFLSHLSQMYNGWVG 105  
Db 299 TTTITLPTNGISATNVIWTDIPAGTTFIPNSVINGVQSNINP----- 344  
QY 106 NMRVIRILAGNAFSAKIIIVCCVPFGTSSSLTIAQATLFPH--VIADVR-----TLEPIE 159

345 -----AGGIQVGTINAGSTTTTTFQVQVTSLPQNGVIRNIGNTTTFYQD-- 388  
QY 160 MPLEDRVNLVYHTNDNQ-PMRLVCMLYPLRTGGSG-----NSDSFVVAG 205  
Db 389 ---DPTKPTITTTNPTPTTVPINTAIINPIKADKTAVIDGDIITYITITFNNDGTVPAT 445  
QY 206 RVLTAPSDDFSFLFL-----VP---PTI-----EOKTRAFTVFNIPLOTL 242  
Db 446 NVIFTDSIPAGTTFIPNSVVLNNPNSPALGITVGLNPGETKTLISFQV----- 497  
QY 243 SNRSRFPSLIQGMILSPDASQVVOFQNGRCLIDGQLLGTTPATSGOLFVRGKINQAGRTL 302  
Db 498 ---RVTOIPAGGTITNEASTTYTQPDPTLPVPTTTEPTPTS-----VTVNTATV 545  
QY 303 NLTEVDGKPFMAFDPSPVGEFDFGKCDWHMIRISKTPNNTSSGDPMSRSVSQVTVNGFVFP 362  
Db 546 NPTK-----SADRAFADIGDIITY-----TISLQNN--GTVP 575  
QY 363 HLGSIQFDEVFNHPTGDIYIG---TIEWISQSPSTPGTDINLM-----E 402  
Db 576 ATNILLTDPIPNGT--FIPNSVTINGISQPNTPSTGITVGLDPTPEAATISFQVQVIS 633  
QY 403 IPDYGSSLSQA-----ANLAPPVFPFGFGEALVYFVSAPFPNNRSAPNDV---PCILLPOE 455  
Db 634 VPPHGLVENQGTVSFTHIVNPEPP-----VTKTSPTKTEAVNTIISTP----- 679  
QY 456 YITHFVSEQAPTMTGDAALLHYVDPDTNRNLGEPKLYPGGYLTCVPGVAGGPOQLPLNGV 515  
Db 680 --TKTADKQLADIGDT--ITYT--ITFRNGTVPATNVTLIDSTPSGTTTIPDSVTINGV 733

RESULT 7  
US-10-052-648A-39  
; Sequence 39, Application US/10052648A  
; Publication No. US2004000558A1  
; GENERAL INFORMATION:  
; APPLICANT: Anderson, David  
; APPLICANT: Burgess, Catherine  
; APPLICANT: Casman, Stacie  
; APPLICANT: Colman, Steven  
; APPLICANT: Edinger, Shlomit R.  
; APPLICANT: Ellerman, Karen  
; APPLICANT: Gunther, Valerie  
; APPLICANT: Kekuda, Ramesh  
; APPLICANT: MacDougall, John R.  
; APPLICANT: Mehraban, Fuad  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Rothenberg, Mark  
; APPLICANT: Shinkets, Richard  
; APPLICANT: Smithson, Glenn  
; APPLICANT: Spyttek, Kimberly A.  
; APPLICANT: Stone, David J.  
; APPLICANT: Vernet, Corine A.M.  
; APPLICANT: Zerhusen, Bryan D.  
; TITLE OF INVENTION: PROTEINS, POLYNUCLEOTIDES ENCODING THEM AND METHODS OF  
; FILE REFERENCE: 21402-250 (CURA-550)  
; CURRENT APPLICATION NUMBER: US/10/052,648A  
; CURRENT FILING DATE: 2002-12-09  
; PRIOR APPLICATION NUMBER: 60/262,454  
; PRIOR FILING DATE: 2001-01-18  
; PRIOR APPLICATION NUMBER: 60/272,920  
; PRIOR FILING DATE: 2001-03-02  
; PRIOR APPLICATION NUMBER: 60/284,549  
; PRIOR FILING DATE: 2001-04-18  
; PRIOR APPLICATION NUMBER: 60/303,229  
; PRIOR FILING DATE: 2001-07-05  
; PRIOR APPLICATION NUMBER: 60/262,892  
; PRIOR FILING DATE: 2001-01-19  
; PRIOR APPLICATION NUMBER: 60/263,605  
; PRIOR FILING DATE: 2001-01-23

; PRIOR APPLICATION NUMBER: 60/269,098  
; PRIOR FILING DATE: 2001-02-15  
; PRIOR APPLICATION NUMBER: 60/264,159  
; PRIOR FILING DATE: 2001-01-25  
; PRIOR APPLICATION NUMBER: 60/265,517  
; PRIOR FILING DATE: 2001-01-31  
; PRIOR APPLICATION NUMBER: 60/271,855  
; PRIOR FILING DATE: 2001-02-27  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 97  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 39  
; LENGTH: 2126  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-052-648A-39

Query Match 4.4%; Score 127.5; DB 15; Length 2126;  
Best Local Similarity 20.9%; Pred. No. 0.07;  
Matches 135; Conservative 63; Mismatches 246; Indels 201; Gaps 30;  
QY 3 MASKAPQADGASGAGQLVPEVNTADP-----LPME---PVAGPTTA 42  
Db 937 MLESDVSEPEGDQGLSLOKDDYGFQSGKLEGEFKQPIAVSSMPQOIGVPTSLTQV 996  
QY 43 VATAGVNMIDWVNNFVQSQGEFTISPNTPGDILFDLQGLHNLNPFSLHLSQMYNG 102  
Db 997 VHSAGR-----RFIVSPVESRLRESKIFTSEIPDPVAASTSQGPGMN--LSHSASS-----1046  
QY 103 WGVNMRVRIILAGNAFSA--GKII--VCCVPPGFTSSSLTIAQATLPHPHVIAVDTLE- 156  
Db 1047 -----LSLQQAPSELKHQGTGPNAPFNHNPPTFSP---FLTSIAGVQTAA 1094  
QY 157 -----PI-EMPLEDVNRVLVHTNDNQPTMRLVCMLYPLRTGGSGNSDSFWAGRYL 208  
Db 1095 STPSVSVPTISPLNDISVWQSEGALEPTDKGI-----GGVITSTGVASGGIT 1144  
QY 209 TAPSSDFSLFLVPPTIEOKTRAFVNPINPLQTLNSRFPSPILQGMILSPDASQVVQFON 268  
Db 1145 TILSVSE-----TPTLSSAVSSSTAPAV--VTVSTTSQP--VQAFSTSGSIASSTGSFPS 1193  
QY 269 GRCLIDGQLLGYTPATSGQLFVRGKINGARTLNLTEDVGK-----PF 312  
Db 1194 -----GTFSTTGTTSVSAVPNAKPTVLLQQVAGNTAGVAIVTSVSTTTPFPA 1243  
QY 313 MAFDGPAPVG-----PFDGK---CDWHMRIKTPNNTSSG-----DPMRSVSVOITNVQ 359  
Db 1244 MASQSLPLGSSATSAPTAEVTVVSNH-SLDKASHSTAGLGLFCAPSSSSSGTAVSS 1302  
QY 360 FVPHLGSIQDFVFNHP--TGDIYGTIEWISQFSTPPGTDINLWEIPDYGSSLSQAANLA 417  
Db 1303 SVSQPGIV-----HPLVISAIASTPVLQPAVPTSTPL-----LPQVFNIP 1344  
QY 418 PVFPP-----PGRGEALVYFVSAPFGNNRSAPNDVPCLLPQBYITHFVSEQAFTMGDA-- 471  
Db 1345 PLVQPVANVAVQOITLIH-----SQPQ--PALLPNQPHTCPEMDADTOSKAPG 1391  
QY 472 -----ALLHYVDDP--TNRNLGEFKLYPGGYLT 497  
Db 1392 IDDIKTELEKRLSRFSEHSSGTQHASVLEPLVVEVTPGIPITTAVAPSKLMTSTTST 1451  
QY 498 CYPNGVGAGPOQLPLNGVFLFVSWSRFYQLKPVGT-----ASTA 537  
Db 1452 CLP-----PTNPLGTAGMPVMPVGTGQGVSTPGTHASAPASTA 1490

## RESULT 8

US-10-369-493-16733  
; Sequence 16733, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.

## RESULT 9

US-10-146-473-50  
; Sequence 50, Application US/10146473  
; Publication No. US20030108888A1  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew  
; APPLICANT: Gout, Ivan

; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 16733  
; LENGTH: 1168  
; TYPE: PRT  
; ORGANISM: Bacillus thuringiensis  
US-10-369-493-16733

Query Match 4.3%; Score 123.5; DB 15; Length 1168;  
Best Local Similarity 18.9%; Pred. No. 0.066;  
Matches 120; Conservative 96; Mismatches 249; Indels 169; Gaps 27;  
QY 3 MASKAPQADGASGAGQLVPEVNTADPMPPEPVAGPTTAVATAGQVN--MIDPMT--- 56  
Db 91 LASSIPQ-----GAIINQALTSYTYIVDPSPQPPVTATSSNTVNTAVVDASLSVIK 142  
QY 57 -VNNFVQSPQGEFTI-----SPNTPGDILFDLQGLHNLNPFSLHLSQMYNGVGNMVR 110  
Db 143 STDLSLVQSDGTITVTVVQVNGNTTANTVTLTDLVPE-----180  
QY 111 ILLAGNAFSAKIIIVCCVPPGFTSSSLTIAQATLPHPHVIAVDTLEPIEMP-----LE 163  
Db 181 -----GTAP-----IPNSVTISGVSVPFAD--PNVGIPLNSITPSEIVTVFQVIVQ 225  
QY 164 DVNRVLVHTNDNQPTMRLVCMLYPLRTGGSGSN-----SDSFVAGVLTAPSSDFSF 217  
Db 226 SIPSTVQSTNARIDYTFIADPTAFIISRTITSNPAFTQISDATILSLKAVNAPQTTGD 285  
QY 218 LFLVPTTIEOKTRAFVNPINPLQTLNSRFPSPILQGMILSPDASQVVQFONGRCLIDG-O 276  
Db 286 ILTYITITLE-----NTGNIPATNL-----IPSDTIPQDTPFVENSFTLNGTA 327  
QY 277 LGITTP-----ATSQLFRVRGK-----INQARTLNLTEDVGKPPMAFD 316  
Db 328 ILGANNPVGVTLPALAAATHLISFQILIKDSFRESITNQSNNTTYTIQDFQGPPIET 387  
QY 317 SPAPVGFDPFGKCDWHMRIKTPNNTS--SGDPMRSVSQVTVQGFVPHLGSIQDFEVEN 374  
Db 388 STSNIVITNFVQA--QLITKTSNPTTVDIGTILYISEVKNI-GNVDAINIVFTDSI--- 442  
QY 375 HPTG-----DYGTIEWISQFSTPPGTDINLWEIPDYGSS--LSQAANLAPVFPFPGFE 427  
Db 443 -PAGTTFVPDSV-TINGVLQGVNPENGIPICITIPANSSKTLFQVQTNPPTEIIVNQ 500  
QY 428 -ALVYFVSAPF--GNNRSAPNDVPCLLPQBYITHFVSEQAFTMGDAALLHYVD-PDTHR 483  
Db 501 SSATYQYVSIPTAPPNRSANSNIVTSLQN--ANIIISVKSDITFASIGQIITVNTLQ 558  
QY 484 NLGEFKLYPGGYLTCPVNGV-----GAGPO-----508  
Db 559 NIGTVPANNTVFIDNIPGTFIEDSLAINNVIQGANPENGVLTGTQNETVTVISFQV 618  
QY 509 ---QLPLNGVFLFVSWSRFYQLKPVGTASTARS 539  
Db 619 QLTNIPVGNVNVINISDTSVEYQIDPSSQIIQRRS 652

```

; APPLICANT: Stockert, Elisabeth
; APPLICANT: Gure, Ali
; APPLICANT: Chen, Yao-Tsung
; APPLICANT: Old, Lloyd
; TITLE OF INVENTION: Breast Cancer Antigens
; FILE REFERENCE: L00461/70130(JRV)
; CURRENT APPLICATION NUMBER: US/10/146,473
; CURRENT FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: US 60/291,150
; PRIOR FILING DATE: 2001-05-15
; NUMBER OF SEQ ID NOS: 82
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 50
; LENGTH: 2971
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-146-473-50

Query Match      4.2%; Score 123; DB 14; Length 2971;
Best Local Similarity 21.1%; Pred. No. 0.31;
Matches 122; Conservative 57; Mismatches 206; Indels 192; Gaps 26;

QY 19 GQLVPEVNT---ADPLMEPVAGP-TTAVATAGOVNMIDP--WIVNMFVQSGEGFTIS 71
Db 809 GRTVVVNNPRAPLGPVVRPPGPELSAQTPGCPVQVLPASLMVS---ASPAGPLIP 865
QY 72 PNNTPGDILFDLQGLPHLPFLSHLSQMYNGWVGNMR-----VRIL 112
Db 866 ASRPPGVL---PPLQNSGSLPQVLPSPLGVLGTSRPTTSLKPTPPAPVRLS 920
QY 113 LAGNARSAGKIIVCCVPPGP-----TSSSLTIAQATLPHPHVIADVRLTEPLEMDEV 165
Db 921 PAPPGPSLSLKPLTVPPGPTFPAAATTTTATAT-----TTAVAPTAPQ- 970
QY 166 RNVLVYHTNDNQTM-----RLVCMLYTPLRTGGSGNSDSFVVAGRVLTAPSSDFS 216
Db 971 RLIL--SPDQARLPSEGVVSIQGLASLAQRFPVANAGGS-KPLTFQIQGNKLTLTGAQVR 1027
QY 217 FLFL-----VPTIEQKTRAFV-----PNIPQLTNSRF 247
Db 1028 QLVAGQPRPLQMPPTMNTGVVXIVVRQAPRDLGTLVPVPLAPAPRPPSGGLPAVLNPR- 1086
QY 248 PSLIQGMILSPDASOVQFQNGRCLIDQLLGT--TPATSGQLFRVRGKINQOGARTINLT 305
Db 1087 PTLTPGRLPTP-----TLGTRAMPPTPLVRLKLVHSPSPVSA 1128
QY 306 EVDGKPFMAFDSP--APVGFDFGKCDWHMRISKTPNNTSGDPMRVSQVTVNQGVFPH 363
Db 1129 SAGGAAPLTISPLHVPSLP--GPASSPMPI---PNSSPLASPVSS-TVSVPLSSSLP- 1181
QY 364 LGSIQFDEVFNHTGYGTIEWISQPSPTPGTIDINLWEIPDYGSSLSQAA-----NL 416
Db 1182 -----ISVPTTLPAASAPLTIP-ISAPLTVSASGALLTSV 1217
QY 417 APPVPP-----PGFEALVYFVSAPP 438
Db 1218 TPPLAPVVPAPGPPSLQPSGASPSASALTIGLATAPSLSSSQTPGHPLLLAPTSSHPV 1277
QY 439 PNNRSAPNDVPCLLPQEIY-THFVSEQAPTWGDAAALL 474
Db 1278 LNSTVAPACSPVLVPASALASFPSPAPNPAQASLL 1314

RESULT 10
US-10-282-122A-67793
; Sequence 67793, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari

; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 67793
; LENGTH: 6310
; TYPE: PRT
; ORGANISM: Pseudomonas putida
; US-10-282-122A-67793

Query Match      4.2%; Score 122.5; DB 12; Length 6310;
Best Local Similarity 19.5%; Pred. No. 1.1;
Matches 106; Conservative 63; Mismatches 197; Indels 177; Gaps 21;

QY 15 ASGAQQLVPEVNTADPLMEPV-----AGPTTAVATAGOVNMIDPWVNNFVQSGPGE 67
Db 1776 ASSALQFEPVPTTAPASPSDLFAEDGSSISGTAPAGTRVEVDHANGTLTGTVVAGPDGS 1835
QY 68 FTIS--PNNTPGDILFDLQGLPHLPFLSHLSQMYNGWVGNMRVIRILLAGNAFSAGKIIV 125
Db 1836 FTVTLDPDPAQTNGELL-----DVVAIDGGVSSL-----1863
QY 126 CCVPGFTSSSLTIAQATLPHPHVIADVRLTEPLEMDEVNVLVYHTNDNQTPMLVCMWL 185
Db 1864 ---PQITAPDITAPAA-----PTELVINADGSSVVTGRAEPGSTRVLA--1904
QY 186 YTPLRTGGSGNSDSFVVAGRVLTAPSSDFSFLFVLPVPTIEQKTRAFVVPNIP--LOTLS 243
Db 1905 -----ADGTVLGSVIVGATGSFS--ITLDPQIDGVEVLQVLTATDAAGNASTAS 1951
QY 244 NSRFPSSLIOG-----MILSPDASQV-----VQFQNGRCLIDGQLLGTTPAT 284
Db 1952 NLTPADIDGDTTPPEAPTNLVINPAGSOLTCRGAEGTSVQVRDG-----AGTVVAT 2003
QY 285 SGOLFVRGKINQO-----ARTLNLTVDGKPFVA-----FDSAPVGFDFGKCDWHM 333
Db 2004 -----GTVPDGTFAITLPAITDGTSTLQVTLTDAAGNVQSGSVATPD-----L 2048
QY 334 RISKTNN-----TSSGDDPMRVSQVTVNQGVFPHLGSIQFDEVFNHTPGDYGTI 384
Db 2049 LAPQPTFELALIDGVTFTGRGEGCATVQVR-DATGSLIGTLVNE-----GTF 2096
QY 385 EWISQPSPTPPGTDINLWEIPDYGSSLSQAANLAPPVFP-----GFGEA 428
```

```
Db 2097 SVTLFPAQANGALDIRVVDAGNSSTPLAFTAPDVTTPAAVSNVVVGADGLVLSGRGEA 2156
QY 429 LV-----YFSAFFPCPNRSAPNDVPCILLPOEYITHFVSEQAPTMGDAALHYYVD 478
Db 2157 GATVQVRDANGSVIGTATVGANGTFLVDLTPAAQFGEQLSLVQTDASGNASEA--LOYEI 2214
QY 479 PDT 481
Db 2215 PAT 2217

RESULT 11
US-10-282-122A-48055
; Sequence 48055, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48055
; LENGTH: 1765
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-48055

Query Match 4.0%; Score 115.5; DB 12; Length 1765;
Best Local Similarity 21.0%; Pred. No. 0.72;
Matches 118; Conservative 61; Mismatches 203; Indels 179; Gaps 27;

QY 7 DAPQSDAGSAGQLVPEVNTADPLPMEPVAGPTT--AVATAGQVNMIDPW---IVNVFV 61
Db 870 DVPSVAVTGNAGAVSVIRS-----KAVLGVLTLPAASVAVTVRLCAPSPKPVGVNV 923
QY 62 QSPQGEFTISPNTPGDIILFDLQGLPHLNPFLSHLSQMTYNGVGNVRVILLAGNAFSA 121
Db 924 QPFDGSVAVVPPNVN-----VPSYTLTVL----- 946
```

```
QY 122 KIIVCCVPFGFTSSSLTTIAQATLPHVFIADVRLTLP-----IEMPLEDVRNVLYH---TND 174
Db 947 -----PASAVPLNVGVASSVLPPEMIAP--TTEPTSSVTVPMTGAPGAVVSPVTVTEA 997
QY 175 NQPTMRLVCMYLTPLRTGGSGNSDSFVVA--GRVLT-----APSSDFSLFLVLP---PT 224
Db 998 DGFVLPAASVAVTPTITVPAGSGVEGVYVHAPLGSAVTVASGVSPFVSMLTVPASAVPD 1057
QY 225 IEQKTRAFPTVPNIPLQTLNSRFPFLIOQMILSPDASQVVFQNGRCLIDGOLLG--TTP 282
Db 1058 SDVPSVAVTVGN-----AGAVVSVIYSSKAV-----LGTLTLP 1089
QY 283 ATSGQLFRVKGKINGKARTLNTEVDGKPFMAFDSPAPVGVFPDFGKCDHMRISKTNN- 341
Db 1090 AVS-----VAVTVRLCAPSPKPVGVN---VQFPDGSVAV-----VVPPNV 1126
QY 342 ----TSSGDPMSRSVSQTNVQGFV--PHLGSIQFDEVFNHPTGDIYGTIEMISQBST-PP 394
Db 1127 VPSYTLTVLPASAVPLNVGVASSVLPPEMIA-----PTEPTSS---VTVPMTGAP 1174
QY 395 GTDINLWEIPDYGSSLSQAANLA--PPVFPFGFGEALVY-----FVS 434
Db 1175 GAVVSPVTVTEADGPVLPAAVAVTPTITVPAGSGVEGVYVHAPLGSAVTVASGVSPFV 1234
QY 435 APP-GPNNRSAPNDVPCLL-----POEYITHFVSEQA-----PTMGDAALHYYVDPTN 482
Db 1235 MLTVAPASAVPDNDVPSVAVTVAGIAGAVVSVIRSKAVLGTLTLPASVAVTVRLCAPSP 1294
QY 483 RNLGEPKLYPGGYLTCVNGV 503
Db 1295 AVVGVNVQFPDGSVAVVPSNV 1315

RESULT 12
US-10-282-122A-46577
; Sequence 46577, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
```

PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 46577  
LENGTH: 1194  
TYPE: PR1  
ORGANISM: Bacillus anthracis  
US-10-282-122A-46577

Query Match 4.0%; Score 114.5; DB 12; Length 1194;  
Best Local Similarity 19.6%; Pred. No. 0.49;  
Matches 87; Conservative 55; Mismatches 154; Indels 147; Gaps 21;  
QY 20 QLVPEVNTADPLPMEBVGATATAGQVNMIDPWVNFVQSPQGEFTISPNN-----74  
DB 555 QNVGNINATDVITDTPAGTIFPNS-----VTINGVSSPCA---NPNSGVNVG 601  
QY 75 --TPGDILFDLQGLPHNPFSLHLSQMYNGWGNMRVILLAGNAPSAGKIIIVCCVPP-G 131  
DB 602 TVTPGQI-----VTLY-----QVTVTALPPDG 624  
QY 132 FTSSSTIIAQ-----ATLPHVIADVLTLEPIEMP-----LEDVRNVLXHTN 173  
DB 625 IIKNTATVTVTQFNPCEPPIITDPTTVEVSIVITPTPNPKLADKQIVDINEIITYTV 584  
QY 174 DNPOTMRLVCMVLTPLRTGGSGNSDSFVVGAGRLVLTAPSDSFSLFLVPPTI-----EQKT 229  
DB 685 TFQ-----NRGVPATSVIIVTDPLA--NGLTFVPGTVILNGIPDL 722  
QY 230 RAFTVPNPIQTLNSRFPSLIOGMILSPDASQVQFONGRCLIDQGLLGTTPATSGQLF 289  
DB 723 GANPVEGIPVGT-----VNPNDTIVQFQ-----ARVTSVPPGGII 758  
QY 290 RVRKINGARTLNLTEVDGKPFMAFDSAPVGFPGKCDMHRISKTPNNTSSGDPMR 349  
DB 759 R-----NOATVTFYEPIDGEPVITDPTPTINTDVTALLNPQKATPETVTLGDIIT 813  
QY 350 -SVSVQTNVQGVFPHLSIQFDEVNHTG-DYTG-----TIEWISQSPSTPGTDINLWEIP 404  
DB 814 YTISLQNT--GTIPANNILVSDPI---PTGTSFIONSVTINNVSQPTANPETGI---QIP 865  
QY 405 DYSSLSQAAN---LAPVFPFG 424  
DB 866 TLPSPSESATISFHLVTSIPPSG 888

RESULT 13  
US-10-327-481A-38  
Sequence 38, Application US/10327481A  
Publication No. US20040001864A1  
GENERAL INFORMATION:  
APPLICANT: King, Andrew M. Q.  
APPLICANT: Burman, Alison J.  
APPLICANT: Audonnet, Jean-Christophe F.  
APPLICANT: Lombard, Michel F.A.  
TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease  
FILE REFERENCE: 454313-3178  
CURRENT APPLICATION NUMBER: US/10/327,481A  
CURRENT FILING DATE: 2002-12-20  
PRIOR APPLICATION NUMBER: PCT/FR01/02042  
PRIOR FILING DATE: 2001-06-27  
PRIOR APPLICATION NUMBER: FR 0008437  
PRIOR FILING DATE: 2000-06-29  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 38  
LENGTH: 1147  
TYPE: PR1  
ORGANISM: Foot-and-mouth disease virus  
US-10-327-481A-38

Query Match 3.9%; Score 113.5; DB 15; Length 1147;  
Best Local Similarity 21.2%; Pred. No. 0.57;  
Matches 109; Conservative 45; Mismatches 188; Indels 171; Gaps 23;  
QY 36 VAGPTTAVATAGQVNMIDPWVNFVQSPQGEFTISPNNTGDIILFDLQGLHNPFLSH 95  
DB 129 VAGPNTSGLETRVQ-----AERFFKFLFDWT---TDKPGYLTUKLELPTDHHGVFGH 179  
QY 96 LSONY-----NGWGNMRVILLAGNAPSAGKIIIVCCVPPGFTSSSLTTIAQATFLPHVIAD 151  
DB 180 LVDSYAMRNGW-----DVEVSAGVQFNGCCLLVAMVPEWKAFTREKYQLTLFPHQFTS 235  
QY 152 VRT--LEPIEMPLEDRNVLXHTNDQPTMRVLVCMVLTPLRTGGSGNSDSFVVGAGRLV 209  
DB 236 PRINMTAHITVPYLVGNR--YDQYKHKPWTLLVWVLSPL-----273  
QY 210 APSDSFSLFLVLPPTTEQKTRAFVTNPIQLTLSNSRFPPL-----IQGMILS 257  
DB 274 -----TVSNTAAPQIKVYANIAITYVHVAGELPS 302  
QY 258 PDASQVQFONGRCLIDGQLLGTTPATSGQLFRVRGKI-----NOGARTLNLTVEV--- 307  
DB 303 KEGIFPVACADGY-----GGLVTTDPKTAADPVY---GKVYNPPKTYNPGRTNLLDVAEAC 355  
QY 308 -----DGKPFM---AFDSPAPVFPDQKCDMHRISKTPNNTSSGDMRMSVSVQTNV 357  
DB 356 PTFLEFDDGKPYVVTADDTL-----LAKFVDSLAAKHMSNTYLSG-----IA 399  
QY 358 QGFVPHLGSQFQDFVFNHPTGDYIGTIEWISQP--STPPGT-----DINL--- 400  
DB 400 QYTYQYSGTINLHFMFTGSDSKARYVAVIPQVETFPDTPBEAAHCHIAEWDTGLNSK 459  
QY 401 --WEIP-----DYSSLSQAANLAPVFPFGFGEALVYFVSAPFGPN-----NRSAPNDV 448  
DB 460 FTFSIPVYSAADYATASDTAETTNVQ-----GWVCYQITHGKAENDTLLVSASAGKDF 514  
QY 449 PCLLPQEVITHFVSEQAPTMGDAALLHVDPDT 481  
DB 515 ELRLPIDPRT-----QTTTGESA-----DEVT 537

RESULT 14  
US-10-224-999A-3480  
Sequence 3480, Application US/10224999A  
Publication No. US20030171318A1  
GENERAL INFORMATION:  
APPLICANT: Myriad Genetics, Inc.  
APPLICANT: Morham, Scott  
APPLICANT: Zavitz, Kenton  
APPLICANT: Hobden, Adrian  
TITLE OF INVENTION: Composition and Method for Treating Viral Infection  
FILE REFERENCE: 5004.01  
CURRENT APPLICATION NUMBER: US/10/224,999A  
CURRENT FILING DATE: 2003-03-03  
PRIOR APPLICATION NUMBER: US 60/313,695  
PRIOR FILING DATE: 2001-08-20  
NUMBER OF SEQ ID NOS: 3484  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 3480  
LENGTH: 1046  
TYPE: PR1  
ORGANISM: Human herpesvirus 5  
US-10-224-999A-3480

Query Match 3.8%; Score 111; DB 14; Length 1046;  
Best Local Similarity 23.2%; Pred. No. 0.85;  
Matches 109; Conservative 50; Mismatches 181; Indels 130; Gaps 28;  
QY 17 GAGQLPVEVNTA---DPLPMEPVAGPT-----TAVATAGQVNMIDPWVNFVQSPQGEF 68  
DB 580 GAGVNPAGAGAAIITFTFVNPSTAPAPATPTTFACTQTPVNGNSPW-----APTAPL 632  
QY 69 --TISPNNTPGDIILFDLQGLPHL--NPFLSHLSQMYNGWGNMRVILLAGNAPSAGKII 124

Db 633 PGDMNPANPRERAWALK-NPHLAYNPP-----RMPTTSTTSONNVS----- 673  
Qy 125 VCCVPPGFTSSSLITIAQATLPHVIAVRL--EPIEMPLEDVRNVLVHTNDONPTMLRV 182  
Db 674 ---TPRRRSTPRAAVTQTASQNADEVWALRDQTAESVEDSEE-----EDDD----- 719  
Qy 183 CMLYPLRTGCGSGNSDSFVAVAGRVLTAPSSDFSLFLVPP--TIEQKTRAFVTP----- 235  
Db 720 -----SSDTGSVVSLSGH--TTPSSDYNDV-ISPSSQTPEQST-----PSRIRK 759  
Qy 236 ---NIPLOTLSNRPSLSLQCMILSPDAS--QVVFQFQNGRCLIDCOLLG---TTPATSG 286  
Db 760 AKLSSPMWTTTSQXK-VLGKXVATPHASARAQTVTSTPVQGRVEKQVSGTSTVPATLL 818  
Qy 287 QLFVRVRGK-----INQARTLNLEVDGKPFMAFDSP-----APVGPDPFGKCDWHRM 334  
Db 819 QQPASSKTTSSRNVTSRGARTSASARQPSASASVLSPTEDDVVSPVTSF-----LSMLSS 874  
Qy 335 ISKTNTNTSSGDMRSVSQVTVNVOGFVPHLGSIOFDEVFNHP-----TGDIYIGTIEWISQ 389  
Db 875 ASPSPAKSAPPSPVKGSRGSRVGPLKPTLGG---KAVVGRPPSPVSGSAPGRLSGTSR 931  
Qy 390 -----PSTPPGTDINLWEIPDYGSSLSQAANLAPPV-----PFPGEAL 429  
Db 932 AASTTPTYPAVTTV----YPPSSTAKSSVSN-APPVAPSILKPGASAAAL 976

RESULT 15

US-10-381-247B-18  
; Sequence 18, Application US/10381247B  
; Publication No. US20040073001A1  
; GENERAL INFORMATION:  
; APPLICANT: Akiyama, Tetsu  
; APPLICANT: Adachi, Shungo  
; APPLICANT: Kyowa Hako Kogyo Co., Ltd.  
; TITLE OF INVENTION: beta-catenin Nuclear Localizing Protein  
; FILE REFERENCE: 082394-000000US  
; CURRENT APPLICATION NUMBER: US/10/381,247B  
; CURRENT FILING DATE: 2003-03-21  
; PRIOR APPLICATION NUMBER: JP 2000-287876  
; PRIOR FILING DATE: 2000-09-22  
; PRIOR APPLICATION NUMBER: WO PCT/JP01/08140  
; PRIOR FILING DATE: 2001-09-19  
; NUMBER OF SEQ ID NOS: 18  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 18  
; LENGTH: 1394  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-381-247B-18

Query Match 3.8%; Score 110.5; DB 12; Length 1394;  
Best Local Similarity 20.2%; Pred. No. 1.5;  
Matches 97; Conservative 41; Mismatches 160; Indels 183; Gaps 25;  
Qy 11 SADGASGAGQLVPEVNTADPLEMEPVAGPTTAVATAGOVNMIDPMVNNFVOSPOGEFTI 70  
Db 1049 SNNNPFMG-----INTQNPRIISGP--NPVVPMTLSPMGMTQPLSHSNQMFSPN---AV 1098  
Qy 71 SPNNTPGDILFDLQGLPHL---NPLSHLSQMYNGWGVNMVRVRIILAGNAFSAGKIIVCC 127  
Db 1099 GPNIPPHGV---PMGPGLSMHNPIMGHGSQ----- 1125  
Qy 128 VPPGFTSSSLITIAQATLPHVIAVRLTLEPIEMPLEDVRNVLVHTN----- 173  
Db 1126 EPFMVPGRMGPFQG--FP-----FVOSPPQV--PFPHNGPSGCGSPFGMGF 1171  
Qy 174 -----DNQPTMRLVCMYLTPLRTGCGSGNSDSFVAVAGRVLTAPSSDFSLFLVPP 224  
Db 1172 PEGPLGRPSNLPOSSADALCKP-----GGPGGDSFTVLGNSMPSVFTD-----PD 1219  
Qy 225 IEQKTR--AFTVENIPLQTLNSRFPFSLIQCMILSPDASQVVQ-FQNGRCLIDGQLLGT 281

Db 1220 LQEVIRPGATGIPBFDLSRIIPSEKP-----SQTLOYFPRGEVPGRKOPQGP 1267  
Qy 282 PATSQQLFRVRGKINQAGARTLNLEVDGKPFMAFDSPAPVGPDPFGKCDWHRMIRISKTPNN 341  
Db 1268 PGFS-----HMQGMGEQAPRMCLA-----LPGMGPGPGVGTPTDI-----PLG 1305  
Qy 342 TSSG-----DPMRSVS--VQTVNVOGFVPHLGSIOFDEVFNHPTGDYIGTIEWISOPSTPPGT 396  
Db 1306 TAPSPMGHNPMPAPPAFLQOGMMG--PHRMMS-----DAQSTMPGQ 1344  
Qy 397 DINLWEIPDYGSSLSQAANLAP-----PVFPFGGALVYFVSAFPGPNR 442  
Db 1345 -----PTLMSNPAAVGMIPGKDRGPAGLYTHFGPVGSPG-----MMMSMQGMGPNRT 1393  
Qy 443 S 443  
Db 1394 S 1394

Search completed: June 1, 2004, 14:04:31  
Job time : 36.9037 secs

**THIS PAGE BLANK (USPTO)**



Result No.	Query			ID	Description
	Score	Match	Length		
1	2887	99.7	546	1	COAT SOUV3
2	1114	38.5	539	1	COAT LORDV
3	328	11.3	2344	1	POLN RHDV
4	299	10.3	703	1	COAT SNSV4
5	293.5	10.1	702	1	COAT SWSV1
6	284	9.8	668	1	COAT FCVF4
7	269.5	9.3	2208	1	POLN MANCV
8	268.5	9.3	671	1	COAT FCVF9
9	267	9.2	668	1	COAT FCVC6
10	153.5	5.3	2193	1	POLG HE71B
11	151.5	5.2	2193	1	POLG HE71M
12	145.5	5.0	2194	1	POLG HE701
13	145	5.0	2193	1	POLG CX16T
14	144	5.0	855	1	POLG HRV3
15	143	4.9	2206	1	POLG POL3L
16	140.5	4.9	2206	1	POLG POL32
17	131.5	4.5	2205	1	POLG POL2M
18	128.5	4.4	2193	1	POLG CX16G
19	128.5	4.4	2207	1	POLG POL2L
20	124.5	4.3	2196	1	POLG EC05N
21	123	4.2	2175	1	POLG BOVEV
22	122.5	4.2	862	1	POLG EC16H
23	122.5	4.2	2209	1	POLG POL1S
24	121.5	4.2	2208	1	POLH POL1M
25	119	4.1	2164	1	POLG HRV89
26	119	4.1	2185	1	POLG SVDVH
27	119	4.1	2185	1	POLG SVDV0
28	119	4.1	2194	1	POLG EC30B
29	118	4.1	832	1	POLG HRV1A
30	117.5	4.1	2206	1	POLG POL1M
31	116.5	4.0	2150	1	POLG HRV2
32	116	4.0	1200	1	HYAL STRPU
33	116	4.0	2185	1	POLG CXBP5

Query Match 38.5%; Score 1114; DB 1; Length 539;  
Best Local Similarity 44.3%; Pred. No. 5.1e-73;



Db 369 RNSAHMGNRWSTISGVSQPRV-----FQSNRHFDFDSTTTGWSTPYYPVBIKIQGK 423  
 Qy 295 INQAGARTLMLTEVDGKPPMAFDSAPVGFPGKCDWHMRISKTPNNTSSGDPNRSYSVQ 354  
 Db 424 VGSNNKWFHVIDT---KALVPGIPDGPDPDTPIDP---ETKATNGNFSYGESYAGST- 475  
 Qy 355 TNVQGFVPHLGSIQDEVPNHTGDYI-GTIEWISQP 390  
 Db 476 -----TIKPNENSTHFKGTYICGTLSTVEIP 501

RESULT 5  
 COAT\_SMSV1 STANDARD; PRT; 702 AA.  
 AC P362B4;  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS San Miguel sea lion virus (serotype 1) (SMSV 1).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesiviruses.  
 NCBI\_TaxID=36406;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92410750; PubMed=1529644;  
 RA "Nucleotide sequence of the capsid protein gene of two serotypes of  
 RT San Miguel sea lion virus: identification of conserved and non-  
 RT conserved amino acid sequences among calicivirus capsid proteins.";  
 RL Virus Res. 24:211-222(1992).  
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC EMBL; M87481; AAA16217.1; -.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DEE3D7 CRC64;

Query Match 10.1%; Score 293.5; DB 1; Length 702;  
 Best Local Similarity 28.3%; Pred. No. 1.3e-13;  
 Matches 102; Conservative 45; Mismatches 178; Indels 35; Gaps 12;  
 Qy 12 ADGAGAGOLVPEVNTADPLNMEPVAGPTTAVATAGQVNMID-PWVNNFVQSGEFTI 70  
 Db 153 SDGPGGADIVTEEGTVQQQVPQAQSAALTLAAASTGKTVDCEWTT---FFSHYTVNW 209  
 Qy 71 SPNPTFGDILFDLQGLPHNLPFLSHLSOMYNGWGNMVRILLAGNAFSAKIIIVCCVPP 130  
 Db 210 STTEAGKTLFRAUSPELNPYLRHLSISYTSWGGIDVRFVSGGVGKLAALIVPP 269  
 Qy 131 GFTS-SSLTIAQATLPPHVIADVTLEPIEMPLEDVNNVLYHTNDNQPTMLVCMYLP 189  
 Db 270 GIEPVESPTMLQ---YPHVLFDAQVPEVIFTIPDIRKTLXHSMDTDTTLRLVMVYNEL 326  
 Qy 190 RTGGGSGNDSFVACRVLTAPSSDFSLFVPPTEIEQKTRAFVTPNIPLOT--LSNSRF 247  
 Db 327 INPYEQSEPKS-SCSITVETRPSSDFTSLLKPPGSLKXGIPSDLIIPNRSRHWGNRW 385

Qy 248 PSLIQMILSPASQVQFQNGRCL-IDQLLGTTPATSGOLFVRVGRKINQOARTLNLTE 306  
 Db 386 WSTIDGFFVVQPRV-----FQSNRHFDFDSTTTGWSTPYYPVBIKIQGK 440  
 Qy 307 VDGKPMFDSAPVGFPGKCDWHMRISKTPNNTSSGDPNRSYSVQTNVQGF 360  
 Db 441 TE-----KSLVP-GLPD-----GWPDITPTAMTASNGNDYTVAEYRIITNNGTHFKGF 488

RESULT 6  
 COAT\_FCVF4 STANDARD; PRT; 668 AA.  
 AC P27405;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS Feline calicivirus (strain Japanese F4) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesiviruses.  
 NCBI\_TaxID=11980;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91306470; PubMed=1853578;  
 RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,  
 RA Miyamura K., Yamazaki S., Mikami T.;  
 RT "Sequence analysis of the 3'-end of feline calicivirus genome.";  
 RL Virology 183:810-814(1991).  
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC EMBL; D90357; BAA14371.1; -.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 668 AA; 73589 MW; 85BDDCB85804E503 CRC64;

Query Match 9.8%; Score 284; DB 1; Length 668;  
 Best Local Similarity 24.7%; Pred. No. 5.8e-13;  
 Matches 145; Conservative 74; Mismatches 246; Indels 122; Gaps 27;  
 Qy 23 PEVNTA-DPLNMEPVAGPTTAVATAGQVNMIDPWVNNFVQSGEFTI-----SPNNT 75  
 Db 133 PEQGTAVGVGVIAPSAQMSAADMASGKSVDSW-----EAFSFTHSVNWSTSET 183  
 Qy 76 PGDILFDLQGLPHNLPFLSHLSOMYNGWGNMVRILLAGNAFSAKIIIVCCVPPGFTS- 134  
 Db 184 QGKILFKQSLGFLNLPYLRHLSISYTSWGGIDVRFVSGGVGKLAALIVPPGVDPV 243  
 Qy 135 SSUTIAQATLPPHVIADVTLEPIEMPLEDVNNVLYHTNDNQPTMLVCMYLP-RTGG 193  
 Db 244 QSTSMQLQ---YPHVLFDAQVPEVIFTIPDLRSTLYHVNSTDTTSLVIMVNDLINPYA 300  
 Qy 194 GSGNDSFVACRVLTAPSSDFSLFVPPTEIEQKTRAFVTPNIPLOT-----R 246  
 Db 301 NDSNSGCVIT---VETKPGDPFKHLLKPP-----GSVLTHGSIPLPKSSSLWIGNR 353



CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M86379; AAA79327.1; --  
CC EMBL; Z11536; CAA77636.1; --  
CC PIR; B43382; VCMWFF9.  
CC InterPro; IPR004005; Calici\_coat.  
CC InterPro; IPR008975; Viral\_Cap\_coat.  
CC Pfam; PF00915; Calici\_coat; 1.  
CC Coat protein; Glycoprotein.  
CC CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC FT CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC SEQUENCE 671 AA; 73441 MW; 33BEE86D8370D5E5 CRC64;  
CC -----

Query Match 9.3%; Score 268.5; DB 1; Length 671;  
Best Local Similarity 26.8%; Pred. No. 7.7e-12;  
Matches 110; Conservative 47; Mismatches 149; Indels 105; Gaps 18;  
10 QADGAGAGQGVPEVNT-ADPLEMEFVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEF 68  
124 EADGSGITA-----PEQCTWVGVIAPFSAQMSTAADMATCKSVDSW-----EAF 170

69 TI-----SPNTPGDILDLQGLPHLPFLSHLSQVNGVGNMVRILLAGNAPSACK 122  
COAT FCVC6 STANDARD; PRT; 668 AA.  
AC P27404;  
DT 01-AUG-1992 (Rel. 23, Created)  
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2  
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein  
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Vesivirus.  
OC NCBI\_TaxID=11979;  
OC [1]  
RN SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
RX MEDLINE=91374597; PubMed=1716692;  
RA Neill J.D., Reardon I.M., Heinrichson R.L.;  
RT "Nucleotide sequence and expression of the capsid protein gene of  
RT feline calicivirus."  
RL J. Virol. 65:5440-5447 (1991).

Query Match 9.3%; Score 267; DB 1; Length 668;  
Best Local Similarity 31.7%; Pred. No. 9.9e-12;  
Matches 76; Conservative 33; Mismatches 97; Indels 34; Gaps 8;  
34 EPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----SPNTPGDILDLQGLP 87  
145 EFNQMSSTAADMATCKSVDSW-----EAFSEHTSVNMTSETQCKILFKQSLGP 195

88 HLNPLSHLSQVNGVGNMVRILLAGNAPSACKIIVCCVPPGFTS-SSLTAAQTLP 146  
196 LLNPVYLTHAKLYVAVSGVDVRFSGVFGKLAALVVPVGPIDPVQSTMLQ----YP 252  
147 HVIADVRLTEPIEMPLEDVRNVLVHTNDNQPTMRLVCMLYTLP-RTGGSGNSDSFVAVG 205  
253 HYLFDARQVPEVIFPDLRSTLYHLSMDTDTTSLVIMVYNDLNPANDSNSSGCIVT- 311

206 RVLTAPSDSFGLFVFPPTIEQKTRAFVTPNIPQLTSLNS-----RFPPLIQGMILSP 258  
312 -VETKPGDPFKHLLKPP-----GSMITHGSIPLSDLPKSSSLWIGNRFSWSDITDFVIRP 365

RESULT 10  
POLG HE71B  
ID POLG HE71B STANDARD; PRT; 2193 AA.  
AC Q66478;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)

Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2  
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein  
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
OS Human enterovirus 71 (strain BrCr) (Ev 71).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Enterovirus.  
OC NCBI\_TaxID=69153;  
OC [1]  
RN SEQUENCE FROM N.A.  
RX STRAIN=BRCR;  
RX MEDLINE=96434998; PubMed=8837884;  
RA Brown B.A., Pallansch M.A.;  
RT "Complete nucleotide sequence of enterovirus 71 is distinct from  
RT

110 QADGAGAGQGVPEVNT-ADPLEMEFVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEF 68  
124 EADGSGITA-----PEQCTWVGVIAPFSAQMSTAADMATCKSVDSW-----EAF 170

69 TI-----SPNTPGDILDLQGLPHLPFLSHLSQVNGVGNMVRILLAGNAPSACK 122  
COAT FCVC6 STANDARD; PRT; 668 AA.  
AC P27404;  
DT 01-AUG-1992 (Rel. 23, Created)  
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2  
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein  
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Vesivirus.  
OC NCBI\_TaxID=11979;  
OC [1]  
RN SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
RX MEDLINE=91374597; PubMed=1716692;  
RA Neill J.D., Reardon I.M., Heinrichson R.L.;  
RT "Nucleotide sequence and expression of the capsid protein gene of  
RT feline calicivirus."  
RL J. Virol. 65:5440-5447 (1991).

Query Match 9.3%; Score 268.5; DB 1; Length 671;  
Best Local Similarity 26.8%; Pred. No. 7.7e-12;  
Matches 110; Conservative 47; Mismatches 149; Indels 105; Gaps 18;  
10 QADGAGAGQGVPEVNT-ADPLEMEFVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEF 68  
124 EADGSGITA-----PEQCTWVGVIAPFSAQMSTAADMATCKSVDSW-----EAF 170

69 TI-----SPNTPGDILDLQGLPHLPFLSHLSQVNGVGNMVRILLAGNAPSACK 122  
COAT FCVC6 STANDARD; PRT; 668 AA.  
AC P27404;  
DT 01-AUG-1992 (Rel. 23, Created)  
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2  
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein  
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Vesivirus.  
OC NCBI\_TaxID=11979;  
OC [1]  
RN SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
RX MEDLINE=91374597; PubMed=1716692;  
RA Neill J.D., Reardon I.M., Heinrichson R.L.;  
RT "Nucleotide sequence and expression of the capsid protein gene of  
RT feline calicivirus."  
RL J. Virol. 65:5440-5447 (1991).

Query Match 9.3%; Score 267; DB 1; Length 668;  
Best Local Similarity 31.7%; Pred. No. 9.9e-12;  
Matches 76; Conservative 33; Mismatches 97; Indels 34; Gaps 8;  
34 EPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----SPNTPGDILDLQGLP 87  
145 EFNQMSSTAADMATCKSVDSW-----EAFSEHTSVNMTSETQCKILFKQSLGP 195

88 HLNPLSHLSQVNGVGNMVRILLAGNAPSACKIIVCCVPPGFTS-SSLTAAQTLP 146  
196 LLNPVYLTHAKLYVAVSGVDVRFSGVFGKLAALVVPVGPIDPVQSTMLQ----YP 252  
147 HVIADVRLTEPIEMPLEDVRNVLVHTNDNQPTMRLVCMLYTLP-RTGGSGNSDSFVAVG 205  
253 HYLFDARQVPEVIFPDLRSTLYHLSMDTDTTSLVIMVYNDLNPANDSNSSGCIVT- 311

206 RVLTAPSDSFGLFVFPPTIEQKTRAFVTPNIPQLTSLNS-----RFPPLIQGMILSP 258  
312 -VETKPGDPFKHLLKPP-----GSMITHGSIPLSDLPKSSSLWIGNRFSWSDITDFVIRP 365

[2]  
RN SEQUENCE FROM N.A.  
RP Neill J.D.;  
RA Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
CC -!- PTM: The N-terminus is blocked.  
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; M32819; AAA2925.1; --  
CC EMBL; U13992; AAC13993.1; --  
CC PIR; A40507; VCMWFF.  
CC InterPro; IPR004005; Calici\_coat.  
CC InterPro; IPR008975; Viral\_Cap\_coat.  
CC Pfam; PF00915; Calici\_coat; 1.  
CC Coat protein; Glycoprotein.  
CC CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
CC SEQUENCE 668 AA; 73550 MW; 9E52312108D441 CRC64;  
CC -----

Query Match 9.2%; Score 267; DB 1; Length 668;  
Best Local Similarity 31.7%; Pred. No. 9.9e-12;  
Matches 76; Conservative 33; Mismatches 97; Indels 34; Gaps 8;  
34 EPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----SPNTPGDILDLQGLP 87  
145 EFNQMSSTAADMATCKSVDSW-----EAFSEHTSVNMTSETQCKILFKQSLGP 195

88 HLNPLSHLSQVNGVGNMVRILLAGNAPSACKIIVCCVPPGFTS-SSLTAAQTLP 146  
196 LLNPVYLTHAKLYVAVSGVDVRFSGVFGKLAALVVPVGPIDPVQSTMLQ----YP 252  
147 HVIADVRLTEPIEMPLEDVRNVLVHTNDNQPTMRLVCMLYTLP-RTGGSGNSDSFVAVG 205  
253 HYLFDARQVPEVIFPDLRSTLYHLSMDTDTTSLVIMVYNDLNPANDSNSSGCIVT- 311

206 RVLTAPSDSFGLFVFPPTIEQKTRAFVTPNIPQLTSLNS-----RFPPLIQGMILSP 258  
312 -VETKPGDPFKHLLKPP-----GSMITHGSIPLSDLPKSSSLWIGNRFSWSDITDFVIRP 365

RESULT 10  
POLG HE71B  
ID POLG HE71B STANDARD; PRT; 2193 AA.  
AC Q66478;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)

Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2  
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein  
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
OS Human enterovirus 71 (strain BrCr) (Ev 71).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Enterovirus.  
OC NCBI\_TaxID=69153;  
OC [1]  
RN SEQUENCE FROM N.A.  
RX STRAIN=BRCR;  
RX MEDLINE=96434998; PubMed=8837884;  
RA Brown B.A., Pallansch M.A.;  
RT "Complete nucleotide sequence of enterovirus 71 is distinct from  
RT

110 QADGAGAGQGVPEVNT-ADPLEMEFVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEF 68  
124 EADGSGITA-----PEQCTWVGVIAPFSAQMSTAADMATCKSVDSW-----EAF 170

69 TI-----SPNTPGDILDLQGLPHLPFLSHLSQVNGVGNMVRILLAGNAPSACK 122  
COAT FCVC6 STANDARD; PRT; 668 AA.  
AC P27404;  
DT 01-AUG-1992 (Rel. 23, Created)  
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2  
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein  
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Vesivirus.  
OC NCBI\_TaxID=11979;  
OC [1]  
RN SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
RX MEDLINE=91374597; PubMed=1716692;  
RA Neill J.D., Reardon I.M., Heinrichson R.L.;  
RT "Nucleotide sequence and expression of the capsid protein gene of  
RT feline calicivirus."  
RL J. Virol. 65:5440-5447 (1991).

Query Match 9.3%; Score 268.5; DB 1; Length 671;  
Best Local Similarity 26.8%; Pred. No. 7.7e-12;  
Matches 110; Conservative 47; Mismatches 149; Indels 105; Gaps 18;  
10 QADGAGAGQGVPEVNT-ADPLEMEFVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEF 68  
124 EADGSGITA-----PEQCTWVGVIAPFSAQMSTAADMATCKSVDSW-----EAF 170

69 TI-----SPNTPGDILDLQGLPHLPFLSHLSQVNGVGNMVRILLAGNAPSACK 122  
COAT FCVC6 STANDARD; PRT; 668 AA.  
AC P27404;  
DT 01-AUG-1992 (Rel. 23, Created)  
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2  
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein  
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Vesivirus.  
OC NCBI\_TaxID=11979;  
OC [1]  
RN SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
RX MEDLINE=91374597; PubMed=1716692;  
RA Neill J.D., Reardon I.M., Heinrichson R.L.;  
RT "Nucleotide sequence and expression of the capsid protein gene of  
RT feline calicivirus."  
RL J. Virol. 65:5440-5447 (1991).

Query Match 9.3%; Score 267; DB 1; Length 668;  
Best Local Similarity 31.7%; Pred. No. 9.9e-12;  
Matches 76; Conservative 33; Mismatches 97; Indels 34; Gaps 8;  
34 EPVAGPTTAVATAGQVNMIDPWVNNFVQSPQGEFTI-----SPNTPGDILDLQGLP 87  
145 EFNQMSSTAADMATCKSVDSW-----EAFSEHTSVNMTSETQCKILFKQSLGP 195

88 HLNPLSHLSQVNGVGNMVRILLAGNAPSACKIIVCCVPPGFTS-SSLTAAQTLP 146  
196 LLNPVYLTHAKLYVAVSGVDVRFSGVFGKLAALVVPVGPIDPVQSTMLQ----YP 252  
147 HVIADVRLTEPIEMPLEDVRNVLVHTNDNQPTMRLVCMLYTLP-RTGGSGNSDSFVAVG 205  
253 HYLFDARQVPEVIFPDLRSTLYHLSMDTDTTSLVIMVYNDLNPANDSNSSGCIVT- 311

206 RVLTAPSDSFGLFVFPPTIEQKTRAFVTPNIPQLTSLNS-----RFPPLIQGMILSP 258  
312 -VETKPGDPFKHLLKPP-----GSMITHGSIPLSDLPKSSSLWIGNRFSWSDITDFVIRP 365

RESULT 10  
POLG HE71B  
ID POLG HE71B STANDARD; PRT; 2193 AA.  
AC Q66478;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)







CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; D00820; BAA1891.1; -  
 CC PIR; A36253; GNNY87.  
 CC HSP; P03300; IPOV.  
 CC MEROPS; C03.UPA; -  
 CC MEROPS; C03.UPB; -  
 CC InterPro; IPR003593; AAA\_ATPase.  
 CC InterPro; IPR004004; Calici\_pol\_hel.  
 CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
 CC InterPro; IPR000199; pept\_3C\_picorn.  
 CC InterPro; IPR000081; peptidase\_C3.  
 CC InterPro; IPR003338; Pico\_P1A.  
 CC InterPro; IPR002527; Pico\_P2B.  
 CC InterPro; IPR001676; Rnv.  
 CC InterPro; IPR000605; RNA\_helicase.  
 CC InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 CC InterPro; IPR001205; RNA\_pol\_P3D.  
 CC InterPro; IPR007094; RNA\_pol\_P3vir.  
 CC InterPro; IPR008975; Viral\_cap\_coat.  
 CC Pfam; PF00548; Cys-protease\_3C; 1.  
 CC Pfam; PF02226; Pico\_P1A; 1.  
 CC Pfam; PF00947; Pico\_P2A; 1.  
 CC Pfam; PF01552; Pico\_P2B; 1.  
 CC Pfam; PF00073; Rnv; 3.  
 CC Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 CC Pfam; PF00910; RNA\_helicase; 1.  
 CC PRINTS; PR00918; CALICVIRUS.  
 CC ProDom; PD001125; Cys\_protease\_3C; 1.  
 CC ProDom; PD001306; Pico\_P2A; 1.  
 CC ProDom; PD001274; Pico\_P2B; 1.  
 CC SMART; SM00382; AAA; 1.  
 CC PolyProtein; Coat protein; Core protein; Thiol protease; Myristate;  
 KW RNA-directed RNA polymerase; Hydrolase; RNA polymerase; Myristate;  
 KW Lipoprotein.  
 FT CHAIN 2 69 COAT PROTEIN VP4.  
 FT CHAIN 70 319 COAT PROTEIN VP2.  
 FT CHAIN 320 561 COAT PROTEIN VP1.  
 FT CHAIN 562 871 COAT PROTEIN VP3.  
 FT CHAIN 872 1014 CORE PROTEIN P2A.  
 FT CHAIN 1015 1113 CORE PROTEIN P2B.  
 FT CHAIN 1114 1443 CORE PROTEIN P2C.  
 FT CHAIN 1444 1532 CORE PROTEIN P3A.  
 FT CHAIN 1533 1554 GENOME-LINKED PROTEIN VP3.  
 FT CHAIN 1555 1737 PICORNAIN 3C.  
 FT CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.  
 FT LIPID 2 2 N-myristoyl glycine (in host) (By  
 FT similarity).  
 FT ACT SITE 1701 1701 PROTEASE (POTENTIAL).  
 FT ACT SITE 1715 1715 PROTEASE (POTENTIAL).  
 SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;  
 Query Match 5.0%; Score 145.5; DB 1; Length 2194;  
 Best Local Similarity 27.1%; Pred. No. 0.028;  
 Matches 59; Conservative 29; Mismatches 103; Indels 27; Gaps 9;  
 QY 32 PMEVPAGPTTAVATAGQVN-----MIDPWVNNFVQSPQ-----EFTISPNNTGCDILF 81  
 DB 342 PAFDFSPPTPEHIFGQVHSMLEIVQIESMWINNVNDASGVERLURVQISAQSDMDQLF 401  
 QY 82 ----DLQL-GPHLNPFLSHLSQWYNGVGNMVRILLAGNAFSAKKIIVCCVPPGFTSSS 136

DB 402 NIPLDIQLEGPLRNTLLGNISRYTHWSGSLMTMFCGSEFTTKLICTYPPG-GSSP 460  
 QY 137 LTIAQTATFPFHVIADVRLTEIEMPLEDRVNVLYHTND-----NQPTMRLVCLMYLTPLR 190  
 DB 461 TDRQAMLAHTHVWDFGLQSSITIIIPWISGSHYEMNTDAKAINANVGYVTCFQMTNLV 520  
 QY 191 TGGSGNSDSFVAGRVUTAPSSPSF-LFLVPPPIEQ 227  
 DB 521 APVGA--ADQCIVGMV--AAKDFNLRLMRDSPDIGQ 554  
 RESULT 13  
 POLG\_CX16T  
 ID POLG CX16T STANDARD; PRT; 2193 AA.  
 AC Q9QF31;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein (Contains: Coat protein VP4 (P1A); Coat protein VP2  
 DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein  
 DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
 DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
 DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)).  
 OS Coxsackievirus A16 (strain Taiwan/5079/98).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 CC NCBI\_TaxID=231417;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21427116; PubMed=11536241;  
 RA Yan J.-J., Su I.-J., Chen P.-F., Liu C.-C., Yu C.-K., Wang J.-R.;  
 RT "Complete genome analysis of enterovirus 71 isolated from an outbreak  
 RT in Taiwan and rapid identification of enterovirus 71 and  
 RT coxsackievirus A16 by RT-PCR";  
 RL J. Med. Virol. 65:331-339(2001).  
 CC -1- FUNCTION: It is thought that the P2C protein attaches to vesicular  
 CC membranes and is associated with viral RNA synthesis.  
 CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
 CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
 CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the  
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
 CC substituted for Gln, and Ser or Thr for Gly.  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC (RNA)(N).  
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AF177911; AAD55085.1; -  
 CC HSP; P03300; IPOV.  
 CC InterPro; IPR003593; AAA\_ATPase.  
 CC InterPro; IPR004004; Calici\_pol\_hel.  
 CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
 CC InterPro; IPR000199; pept\_3C\_picorn.  
 CC InterPro; IPR000081; peptidase\_C3.  
 CC InterPro; IPR003338; Pico\_P1A.  
 CC InterPro; IPR002527; Pico\_P2B.  
 CC InterPro; IPR001676; Rnv.  
 CC InterPro; IPR000605; RNA\_helicase.  
 CC InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 CC InterPro; IPR001205; RNA\_pol\_P3D.  
 CC InterPro; IPR007094; RNA\_pol\_P3vir.

```
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00548; Cys-protease-3C; 1.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00947; Pico_P2A; 1.
DR Pfam; PF01552; Pico_P2B; 1.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICVIRUSNS.
DR ProDom; PD001125; Cys_protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 323 COAT PROTEIN VP2.
FT CHAIN 324 565 COAT PROTEIN VP3.
FT CHAIN 566 862 COAT PROTEIN VP1.
FT CHAIN 863 1012 CORE PROTEIN P2A.
FT CHAIN 1013 1111 CORE PROTEIN P2B.
FT CHAIN 1112 1440 CORE PROTEIN P2C.
FT CHAIN 1441 1526 CORE PROTEIN P3A.
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1549 1731 PICORNAIN 3C.
FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.
FT CHAIN 2 2 N-myristoyl glycine (in host) (By
FT LIPID similarity).
FT ACT_SITE 1695 1695 PROTEASE (POTENTIAL).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
FT ACT_SITE 2193 2193 PROTEASE (POTENTIAL).
FT SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;

Query Match 5.0%; Score 145; DB 1; Length 2193;
Best Local Similarity 20.28; Pred. No. 0.031;
Matches 105; Conservative 171; Indels 184; Gaps 22;

QY 20 QLVPEVN---TADPLMEPVAGPTAVATAGQVNMIDPWVNVFQSPQGEFTISNNWP 76
DB 373 ETILEVNNLKTNETPQRLCFVS-----VQSKTGLCAAFRADP 413
QY 77 GDILFDLQGLHNPFLSHLSQYNGVGNVRILLAGNAFSAAGKIIVCCVPPGFT--- 133
DB 414 G-----RDGPWQSTILQGLCRYTQMSGLSEVTFMFAGSGFMATGKMLIATPPGGNVPA 467
QY 134 -----SSSLTI-----AQATLFPHVIADVRTL-----E 156
DB 468 DRITAMLGTHVWDFGLQSSVTLVVPVMSNTHYRAHARAGYDYTTTGIITWIYQTNVYV 527
QY 157 PIEMP-----LEDVR-----NLVHTNDNQPTMRLVC 183
DB 528 PIGAPTAYIVALAAQDNFTMKLCKDTEIQGTANTQGDPIADIMQDVNNQVNSLTA 587
QY 184 MLYTP-----LRTGGSGNSDSFVAGRL-----TAPSS 213
DB 588 LQVLFTADTEASSHRLGTGVVPALQAAETGASSNASDKLIETRCVNLNHHSTQETAIGN 647
QY 214 DFSFLFLV-----PTTIQKTRAFVNPIDLTLSNRP-SLIQGMILSPDASQVQVQ 267
DB 648 FFSRAGLVSIITWPTTGTQNTDGYVNWIDLMGYAQLRKKCELFYMRFDAEFTFVAKP 707
QY 268 NGRCLIDGQL-----GTTTATSGQLFRVGRKLNQARTLNLTVDGKPFMAFDSFAP 320
DB 708 NGELV--PQLLQVMYVPPGAPKPTSRDSPAQTATNPVS-FVKMTDPPAQVSVFPMSPAS 764
QY 321 V-----GFPDFGKCDWHRIS-----KTPNNNTSGDPMRSVSVQTNVQGVFPHLSIQF 369
DB 765 AQWQFYDGYPTFGE---HLQANDLDYGQCPNNMGTFSIRTVGTEKS-----PHSITLRV 816
QY 370 DEVFNHTPDGYGTIEWISQP-----STPP---GTDI 398
DB 817 YMRKIHVRA-----WIPRLRNQPYLFKTPNPKYKNDI 849
```

```
RESULT 14
POLG HRV3 STANDARD; PRT; 855 AA.
ID AC Q82081;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D)] (Fragment).
OS Human rhinovirus 3 (HRV-3).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Rhinovirus.
OX NCBI_TaxID=44130;
XP [1] SEQUENCE FROM N.A., AND X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS).
RP MEDLINE=97094977; PubMed=8939746;
RA Zhao R., Pevear D.C., Kremer M.J., Giranda V.L., Kofron J.A.,
RA Kuhn R.J., Rosemann M.G.;
RT "Human rhinovirus 3 at 3.0-A resolution.";
RL Structure 4:1205-1220(1996).
CC -!- PFM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U60874; AAB05616.1; -.
CC PDB; 1RHI; 12-MAR-97.
DR InterPro; IPR003138; Pico_P1A.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00073; rhv; 3.
KW Polyprotein; Coat protein; Myristate; 3D-structure; Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 331 COAT PROTEIN VP2.
FT CHAIN 332 567 COAT PROTEIN VP3.
FT CHAIN 568 855 COAT PROTEIN VP1.
FT LIPID 2 2 N-myristoyl glycine (in host).
FT HELIX 36 38
FT TURN 50 50
FT TURN 51 54
FT HELIX 57 57
FT STRAND 63 64
FT TURN 83 87
FT STRAND 88 89
FT TURN 90 94
FT STRAND 101 102
FT STRAND 103 105
FT TURN 113 115
FT TURN 123 123
FT STRAND 126 128
FT HELIX 129 129
FT TURN 133 134
FT STRAND 138 140
FT STRAND 142 143
FT TURN 147 151
FT STRAND 152 152
FT TURN 153 155
FT TURN 156 157
FT HELIX 159 165
FT TURN 166 167
FT TURN 168 180
FT TURN 185 186
FT STRAND 188 197
FT TURN 198 198
FT STRAND 203 203
FT TURN 207 208
```

FT	HELIX	213	216	FT	STRAND	686	702
FT	HELIX	219	221	FT	STRAND	714	720
FT	STRAND	223	224	FT	TURN	722	723
FT	TURN	225	226	FT	TURN	730	731
FT	TURN	231	232	FT	HELIX	733	736
FT	STRAND	234	234	FT	STRAND	742	746
FT	TURN	238	244	FT	TURN	747	748
FT	STRAND	245	245	FT	STRAND	750	755
FT	HELIX	247	252	FT	STRAND	764	765
FT	STRAND	255	259	FT	STRAND	770	771
FT	TURN	260	262	FT	STRAND	780	780
FT	STRAND	265	270	FT	HELIX	784	786
FT	STRAND	273	279	FT	STRAND	790	795
FT	STRAND	284	284	FT	STRAND	804	822
FT	STRAND	287	298	FT	STRAND	826	826
FT	TURN	301	302	FT	TURN	833	834
FT	STRAND	307	323	FT	TURN	840	841
FT	TURN	339	342	FT	TURN	852	853
FT	TURN	346	347	FT	NON_TER	855	855
FT	STRAND	354	354	SQ	SEQUENCE	855 AA; 94300 MW; E3F9C92CA2DA8AB1	CRC64;
FT	TURN	357	358	Query Match			
FT	TURN	370	371	Best Local Similarity			
FT	STRAND	373	373	Score 144; DB 1; Length 855;			
FT	STRAND	374	379	Pred. No. 0.01;			
FT	TURN	382	383	Matches 101; Conservative 65; Mismatches 187; Indels 152; Gaps 20;			
FT	STRAND	395	398	QY 2 MMASKDAPOSADGASGAGOLVPEVNTADPLME-----PVAGPTTAVATAGQVMN 51			
FT	STRAND	399	402	Db LMVVEIAPLNA--PTGSSPTLPVTVTIAPMCTEFTGIRSRISIVPOGLPTTTLPGSQFLT 346			
FT	STRAND	404	404	QY 52 ID-----PWVNVFVQSQEFTISPNNT----- 75			
FT	TURN	405	405	Db TDDRQSPSALPSYEPTPRIHIPGKVRNLEIIQVGTLPNMNNTGNDNVNLYLPLHADR 406			
FT	STRAND	410	415	QY 76 PGDILFDLQL----GPHLNPFLSHLSQMYNGVGNMVRILLIAGNAFSAKIIIVCCVPPG 131			
FT	TURN	418	419	Db 407 QNEIQFTKLYIGDGVFKTLLGEIAQYTHWSGSLRSLMYTGPALSSAKIILAYTPPG 466			
FT	HELIX	421	423	QY 132 FTSSSLTIAQATLPHVIADVRLTLEPIEMPLEDRNVLVYHTNDNOFTWRLVCMLYTPLRT 191			
FT	STRAND	424	425	Db 467 -TRGPEDRKKAMLGTHVVVDIGLQSTIVMTIPWTSQVQFR-----YTDPT 511			
FT	TURN	433	434	QY 192 GGGGNSDSFVVAGRVLTAPSSD----PSFLVLPPTTEQKTRFTVNPNIPLQTLNSRF 247			
FT	STRAND	435	439	Db 512 YTSAGYLSCWYLTSLILPPQTSQVYLLSFISACP---DFKLRLMK-----DTQTISQT-- 562			
FT	STRAND	442	448	QY 248 PSLIQMILSPDASQVVOFQNGRCLID-----GOLLGTTPTAT----- 284			
FT	TURN	452	453	Db 563 DALTEG--LSDELEEVIVEKTKQTLASVSSGPKHTQSVPALETANETGATLTPRSDNVET 620			
FT	STRAND	455	455	QY 285 -----SQGLFRVRGKINQAGARTLNLTVEVCGKPFMAFDPSPAPVGFDPFGKCDMMHMRISK 337			
FT	STRAND	457	463	Db 621 RTTYMHFNGSETDVSFLGRAA-CVHVTEIKKNAAAGLDNHRKLEGN----DWKINLSS 675			
FT	TURN	465	466	QY 338 TPNNTSSGDPMRSVSQVNVQGFVPHLGSIQDFEVNHPHTGYIGTIEWISQSPSPPGTD 397			
FT	HELIX	473	476	Db 676 L-----VQLRKLELFT----YVRDSESYT-----ILATASQP----- 704			
FT	TURN	477	478	QY 398 INLWEIPDYGSSLSQAANLAPVFP 422			
FT	STRAND	480	485	Db 705 ----EASSYSNLTQAMVYPPGAP 725			
FT	STRAND	491	496	RESULT 15			
FT	STRAND	505	506	POLG POL3L			
FT	HELIX	511	513	ID POLG POL3L STANDARD; PRT: 2206 AA			
FT	STRAND	517	522	AC P03302; Q84783; Q84784; Q84785; Q84786; Q84787; Q84788; Q84789;			
FT	TURN	526	527	AC Q84790; Q98592; Q98593; Q98594;			
FT	TURN	530	531	DT 21-JUL-1986 (Rel. 01, Created)			
FT	STRAND	536	544	DT 21-JUL-1986 (Rel. 01, Last sequence update)			
FT	TURN	546	547	DT 10-OCT-2003 (Rel. 42, Last annotation update)			
FT	STRAND	549	553	DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins			
FT	STRAND	585	587	DE P2A TO P2C, P3A; Genome-linked protein VPg; Picornain 3C			
FT	STRAND	590	590	DE (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D			
FT	TURN	599	600				
FT	STRAND	601	602				
FT	HELIX	604	606				
FT	TURN	614	616				
FT	TURN	617	617				
FT	STRAND	620	620				
FT	STRAND	623	624				
FT	HELIX	630	632				
FT	STRAND	633	633				
FT	STRAND	634	637				
FT	TURN	638	638				
FT	STRAND	642	651				
FT	TURN	655	656				
FT	HELIX	660	663				
FT	TURN	664	664				
FT	STRAND	666	669				
FT	TURN	677	683				
FT	TURN	684	685				

DE (EC 2.7.7.48)]].

OS Poliovirus type 3 (strains P3/Leon/37 and P3/Leon 12A[11B]).

OC Viruses; sRNA positive-strand viruses, no DNA stage; Picornaviridae;

OC Enterovirus.

OX NCBI\_TaxID=12088;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=P3/Leon/37; PubMed=6324200;

RX MEDLINE=84170338; PubMed=6324200;

RA Stanway G., Hughes P.J., Mountford R.C., Reeve P., Minor P.D.,

RA Schild G.C., Almond J.W.;

RA "Comparison of the complete nucleotide sequences of the genomes of

RT the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin

RT vaccine derivative P3/Leon 12a1b.;"

RL Proc. Natl. Acad. Sci. U.S.A. 81:1539-1543(1984).

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=P3/Leon 12A[11B];

RX MEDLINE=83299239; PubMed=6310508;

RA Stanway G., Cann A.J., Hauptmann R., Hughes P.J., Clarke L.D.,

RA Mountford R.C., Minor P.D., Schild G.C., Almond J.W.;

RA "The nucleotide sequence of poliovirus type 3 Leon 12 alb: comparison

RT with poliovirus type 1.;"

RL Nucleic Acids Res. 11:5629-5643(1983).

RN [3]

RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-878.

RX MEDLINE=95120467; PubMed=7820548;

RA Grant R.A., Hiramath C.N., Filman D.J., Syed R., Andries K.,

RA Hogle J.M.;

RA "Structures of poliovirus complexes with anti-viral drugs:

RT implications for viral stability and drug design.;"

RL Curr. Biol. 4:784-797(1994).

RN [4]

RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1-878.

RA Hiramath C.N., Grant R.A., Filman D.J., Hogle J.M.;

RA "Binding of the antiviral drug wint51711 to the Sabin strain of type-3

RT poliovirus - structural comparison with drug-binding in rhinovirus-

RT 14.;"

RL Acta Crystallogr. D 51:473-489(1995).

CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN

CC O/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.

CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the

CC poliovirus polypeptide. In other picornavirus reactions Gln may be

CC substituted for Gln, and Ser or Thr for Gly.

CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

CC {RNA} (N).

CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,

CC each of which is composed of one copy each of proteins VP1, VP2,

CC VP3, and VP4.

CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC -!- MISCELLANEOUS: THE SEQUENCE OF STRAIN SABIN VACCINE P3/LEON/37 IS

CC SHOWN.

CC -!- MISCELLANEOUS: THE STRAIN SABIN VACCINE P3/LEON/37 IS THE

CC PROGENITOR OF THE STRAIN SABIN VACCINE P3/LEON 12A[11B].

CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.ebi.ac.uk/announcements>

CC or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).

CC -----

CC EMBL; X01392; AAA46914.1; -.

DR EMBL; X00925; CAA25444.1; -.

DR PDB; 1PIV; 03-JUN-95.

DR PDB; 1PVC; 15-SEP-95.

DR PDB; 1VBA; 11-JUL-96.

DR PDB; 1VBB; 11-JUL-96.

DR PDB; 1VBC; 11-JUL-96.

DR PDB; 1VBE; 11-JUL-96.

DR MEROPS; C03.001; -.

DR MEROPS; C03.020; -.

DR InterPro; IPR004004; Calici\_pol\_hel.

DR InterPro; IPR009003; Cys\_ser\_trypsin.

DR InterPro; IPR000199; Pept\_3C\_picorn.

DR InterPro; IPR000081; Peptidase\_C3.

DR InterPro; IPR003138; Pico\_P1A.

DR InterPro; IPR002537; Pico\_P2B.

DR InterPro; IPR001676; Rhv.

DR InterPro; IPR000605; RNA\_helicase.

DR InterPro; IPR007095; RNA\_pol\_DS\_PS.

DR InterPro; IPR001205; RNA\_pol\_P3D.

DR InterPro; IPR007094; RNA\_pol\_PSVir.

DR InterPro; IPR008975; Viral\_cap\_coat.

DR Pfam; PF00548; Cys-protease\_3C; 1.

DR Pfam; PF02226; Pico\_P1A; 1.

DR Pfam; PF00947; Pico\_P2A; 1.

DR Pfam; PF01552; Pico\_P2B; 1.

DR Pfam; PF00073; rhv; 3.

DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.

DR Pfam; PF00910; RNA\_helicase; 1.

DR PRINTS; PR00918; CALICIVIRUS.

DR ProDom; PD001125; Cys\_protease\_3C; 1.

DR ProDom; PD001306; Pico\_P2A; 1.

DR ProDom; PD001274; Pico\_P2B; 1.

DR Polyprotein; Coat protein; Core protein; Transferase;

KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;

KW 3D-structure; Lipoprotein.

FT CHAIN 2 69

FT CHAIN 70 340

FT CHAIN 341 578

FT CHAIN 579 878

FT CHAIN 879 1027

FT CHAIN 1028 1124

FT CHAIN 1125 1453

FT CHAIN 1454 1540

FT CHAIN 1541 1562

FT CHAIN 1563 1745

FT CHAIN 1746 2206

FT LIPID 2

FT ACT SITE 1709 1709

FT ACT SITE 1723 1723

FT VARIANT 431 431

FT VARIANT 864 864

FT VARIANT 908 908

FT STRAND 4 8

FT STRAND 25 29

FT HELIX 36 38

FT TURN 50 50

FT HELIX 51 54

FT STRAND 57 57

FT TURN 63 64

FT TURN 80 81

FT STRAND 83 87

FT TURN 88 89

FT STRAND 90 94

FT STRAND 101 102

FT HELIX 103 105

FT TURN 113 115

FT STRAND 123 123

FT HELIX 126 128

FT TURN 129 129

FT STRAND 133 134

FT STRAND 138 141

FT TURN 142 143

FT STRAND 147 151

FT TURN 152 152

FT HELIX 153 155

FT TURN 156 157

FT HELIX 159 167

FT STRAND 168 180

FT TURN 185 186

FT STRAND 187 197

FT TURN 198 198

COAT PROTEIN VP4.

COAT PROTEIN VP2.

COAT PROTEIN VP1.

CORE PROTEIN P2A.

CORE PROTEIN P2B.

CORE PROTEIN P2C.

CORE PROTEIN P3A.

GENOME-LINKED PROTEIN VP6.

PICORNAIN 3C.

RNA-DIRECTED RNA POLYMERASE P3D.

N-myristoyl glycine (in host).

PROTEASE (POTENTIAL).

PROTEASE (POTENTIAL).

S -> F (IN P3/LEON 12A[11B]).

K -> R (IN P3/LEON 12A[11B]).

T -> A (IN P3/LEON 12A[11B]).

	FT	STRAND	203	203	
	FT	TURN	208	209	
	FT	HELIX	213	216	
	FT	HELIX	219	221	
	FT	STRAND	223	224	
	FT	STRAND	226	226	
	FT	TURN	235	236	
	FT	STRAND	242	242	
	FT	STRAND	245	245	
	FT	HELIX	246	248	
	FT	TURN	249	252	
	FT	HELIX	255	260	
	FT	STRAND	263	267	
	FT	TURN	268	270	
	FT	STRAND	273	278	
	FT	STRAND	287	287	
	FT	TURN	289	291	
	FT	STRAND	292	292	
	FT	STRAND	295	307	
	FT	TURN	308	309	
	FT	STRAND	310	310	
	FT	STRAND	314	331	
	FT	STRAND	337	337	
	FT	TURN	348	351	
	FT	TURN	355	356	
	FT	STRAND	363	363	
	FT	TURN	366	367	
	FT	STRAND	379	380	
	FT	STRAND	382	382	
	FT	TURN	383	383	
	FT	HELIX	384	387	
	FT	TURN	388	388	
	FT	STRAND	391	392	
	FT	STRAND	397	397	
	FT	TURN	399	403	
	FT	HELIX	405	408	
	FT	STRAND	410	413	
	FT	TURN	414	415	
	FT	TURN	418	419	
	FT	STRAND	423	426	
	FT	TURN	429	431	
	FT	TURN	433	437	
	FT	HELIX	439	444	
	FT	TURN	445	446	
	FT	STRAND	447	451	

Query Match                     4.9%; Score 143; DB 1; Length 2206;

Best Local Similarity       21.4%; Pred. No. 0.043;

Matches 123; Conservative   62; Mismatches 220; Indels 170; Gaps   27

Qy	17	GAGQLVPEVNTADLPMEPVAGPTTAVATAGOV-NM-----IDPMIVNNFVSQSGPQ----	66
Dd	349	GSNOYLISDNHQSPCAI-PEDVTTPDIPGEVKRMMLAEIDTWPLNLESTKRNWTMDM	407
Qy	67	-EFTISPNNTPGDILFDLQLGPHLNPFLLSH-----LSQMVGWGVGNMRVRILLAGNAFSA	120
Dd	408	YRVTLSDSADLSQPILCLSLSPASDPRLSHTMTLMGEVLNYTHWAGSLKFTFLFCGSMMAT	467
Qy	121	GKIIVCCVPQGFTSSSLTIQAATLFPHVIADVRTLEPIEMPLEDVNRVLYH--TNDNQPT	178
Dd	468	GKILVAYAPPG-AQPPTSKEAMLGTHVIDWLGLQSSCTMVVPWISNVTYRQTQDSFTE	526
Qy	179	MRLVCMLEY-----NPLRTGGSGNSDSFVVAGRVLTA PSSDFSLFPVPPTEIQKTRAF	233
Dd	527	GGYISMFXQTRIVVPLST-----PKSMSMLGFV--SACNDFSVELLRDTT-----	569
Qy	234	VPIINPLQTLNSRRFPSLTICGMILSPDASQVVVFQNGRCRLIDGQLLGTTTPTSQGLFRVRG	293
Dd	570	-----HISQSALPOGIEDLI-----	585
Qy	294	KINGARTLNLI-TEVDGKPFMAFDSPAPVGPFPDGKCDWMHR--ISKPTNNNTSGDPMR-	349
Dd	586	EVAOGALTSLSPKQODSLPDFTFKASGPA-----HSKEVPALTAVETGATNP LAP	633

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time.31.7889 Seconds  
(without alignments)  
5419.275 Million cell updates/sec

Title: US-09-926-799-3  
Perfect score: 2895  
Sequence: 1 MMASKDAPQSDAGSAGAGQ.....OLKPGVTASTARSLGVRR 546

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL\_25.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriap.\*
  - 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2895	100.0	546	12 Q8JW44	Q8JW44 norwalk-lik
2	2868	99.1	544	12 Q8VA27	Q8VA27 human calic
3	2865	99.0	546	12 Q9IV47	Q9IV47 human calic
4	2231.5	77.1	544	12 Q9DU46	Q9DU46 chiba virus
5	2228.5	77.0	544	12 Q9QTE7	Q9QTE7 chiba virus
6	2215.5	76.5	544	12 Q9I185	Q9I185 human calic
7	2210.5	76.4	544	12 Q9I185	Q9I185 norwalk-lik
8	2208	76.3	543	12 Q9IV43	Q9IV43 human calic
9	2203.5	76.1	544	12 Q9IEJ2	Q9IEJ2 human calic
10	2202	76.1	543	12 Q9I8B9	Q9I8B9 norwalk-lik
11	2201.5	76.0	544	12 Q9I8D1	Q9I8D1 norwalk-lik
12	2199.5	76.0	544	12 Q9IV36	Q9IV36 human calic
13	2179.5	75.3	544	12 Q9IV41	Q9IV41 human calic
14	2178	75.2	543	12 Q8UXJ0	Q8UXJ0 norwalk-lik
15	2159.5	74.6	531	12 Q8V7J5	Q8V7J5 norwalk-lik
16	2070	71.5	540	12 Q9YS14	Q9YS14 norwalk vir

17	2058	71.1	540	12 Q8BCA5	Q8BCA5 human calic
18	2030.5	70.1	530	12 Q838B4	Q838B4 norwalk vir
19	2029.5	70.1	530	12 Q9QT39	Q9QT39 norwalk-lik
20	2003.5	69.2	539	12 Q8JVV5	Q8JVV5 norwalk-lik
21	1993.5	68.9	530	12 Q83876	Q83876 norwalk vir
22	1991.5	68.8	517	12 Q9JH41	Q9JH41 norwalk-lik
23	1986.5	68.6	541	12 Q9IV42	Q9IV42 human calic
24	1985.5	68.6	545	12 Q8VA02	Q8VA02 human calic
25	1981	68.4	544	12 Q66418	Q66418 desert shie
26	1970.5	68.1	543	12 Q8BCA3	Q8BCA3 human calic
27	1964.5	67.9	545	12 Q9I8C8	Q9I8C8 norwalk-lik
28	1959.5	67.7	545	12 Q9WI82	Q9WI82 norwalk-lik
29	1953.5	67.5	545	12 Q9I8C2	Q9I8C2 norwalk-lik
30	1951.5	67.4	545	12 Q9IV45	Q9IV45 human calic
31	1947.5	67.3	539	12 Q9IV48	Q9IV48 human calic
32	1932.5	66.8	545	12 Q9I1H2	Q9I1H2 human calic
33	1298	44.8	522	12 Q8B558	Q8B558 bovine ente
34	1297	44.8	522	12 Q8B4Y9	Q8B4Y9 norwalk-lik
35	1290	44.6	522	12 Q8B4Z2	Q8B4Z2 norwalk-lik
36	1288	44.5	522	12 Q8B4Z5	Q8B4Z5 norwalk-lik
37	1286	44.4	520	12 Q7TBK7	Q7TBK7 bovine ente
38	1286	44.4	522	12 Q8V629	Q8V629 norwalk-lik
39	1285	44.4	522	12 Q8V628	Q8V628 norwalk-lik
40	1282	44.3	520	12 Q7TBL1	Q7TBL1 bovine ente
41	1225	42.3	556	12 Q9PY75	Q9PY75 human calic
42	1225	42.3	556	12 Q9I7V6	Q9I7V6 norwalk-lik
43	1222.5	42.2	519	12 Q9YQ22	Q9YQ22 bovine cali
44	1218	42.1	556	12 Q9I7V9	Q9I7V9 norwalk-lik
45	1185	40.9	535	12 Q9I5C6	Q9I5C6 human calic

## ALIGNMENTS

### RESULT 1

Q8JW44 Q8JW44 PRELIMINARY; PRT; 546 AA.  
 AC Q8JW44;  
 DT 01-OCT-2002 (Tremblrel. 22, Created)  
 DT 01-OCT-2002 (Tremblrel. 22; Last sequence update)  
 DT 01-OCT-2003 (Tremblrel. 25; Last annotation update)  
 DE Capsid protein.  
 GN CAPSID.  
 OS Norwalk-like virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OX NCBI\_TaxID=95340;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Funabashi258;  
 RA Katayama K., Takeda N., Natori K.;  
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Funabashi258;  
 RA Natori K., Takeda N.;  
 RT "Genetic and antigenic relationship among Norwalk-like viruses.";  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AB078335; BAC05516.1;  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 546 AA; 58834 MW; 1965F054E2C481E6 CRC64;

Query Match 100.0%; Score 2895; DB 12; Length 546;  
 Best Local Similarity 100.0%; Pred. No. 3.2e-231;  
 Matches 546; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGTTAVATAGVNMIDPWI VNNF 60  
 Db 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGTTAVATAGVNMIDPWI VNNF 60  
 QY 61 VOSQGEFTISPNTFGDILFDLQLPHLPFLSHLSQWYNGWGNRVRILLAGNAFSA 120

```

Db 61 VKSQPGFTTSSSLTIAQATLFPFHVADVRLTLEPIEMPLEDVNRVLYHTNDNQPTMR 182
121 GKIIIVCCVPPGFTSSSLTIAQATLFPFHVADVRLTLEPIEMPLEDVNRVLYHTNDNQPTMR 180
121 GKIIIVCCVPPGFTSSSLTIAQATLFPFHVADVRLTLEPIEMPLEDVNRVLYHTNDNQPTMR 180
181 LVCMLYPLRTGGGSGNSDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPLQ 240
181 LVCMLYPLRTGGGSGNSDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPLQ 240
241 TLSNSRPSSLIQGMILSPDASQVVFQNGRCLIDGQLGTTTATSGQLFRVRGKINQAR 300
241 TLSNSRPSSLIQGMILSPDASQVVFQNGRCLIDGQLGTTTATSGQLFRVRGKINQAR 300
301 TNLTEVDGKPFMAFDSAPVGFDFGKCDWHMRISKTPTNNTSSGDPMRSVSVQTNVQGF 360
301 TNLTEVDGKPFMAFDSAPVGFDFGKCDWHMRISKTPTNNTSSGDPMRSVSVQTNVQGF 360
361 VPHLGSIQFDEVFNHPTGDIYGTIEWISQPTPGTDLINLWEIPDYGSSLSQAANLAPPV 420
361 VPHLGSIQFDEVFNHPTGDIYGTIEWISQPTPGTDLINLWEIPDYGSSLSQAANLAPPV 420
421 FPGFGEALYVFSAPFPNNRSPNDVPCLLPQEIYTHFVSEQAPTMGDALHLYVDP 480
421 FPGFGEALYVFSAPFPNNRSPNDVPCLLPQEIYTHFVSEQAPTMGDALHLYVDP 480
481 TNRNLGFEKLYPGGYLTCVPGNGVAGPQQLPLNGVLFVSWVSRYQLKPVGTASTARSG 540
481 TNRNLGFEKLYPGGYLTCVPGNGVAGPQQLPLNGVLFVSWVSRYQLKPVGTASTARSG 540
541 LGVRR 546
541 LGVRR 546

```

```

RESULT 2
Q8VA27 PRELIMINARY; PRT; 544 AA.
AC Q8VA27;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/C59/99.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=177069;
RN [1]
RP SEQUENCE FROM N.A.
RA Parkas T., Wilton N., Altaye M., Davies D., Chapman F., Thornton S.A.,
RA Jiang X.;
RT "Homologous vs. heterologous immune responses to Norwalk-like viruses
RT among crew members following acute gastroenteritis outbreaks on two US
RT Navy vessels.";
RL Submitted (0CT-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL; AF435807; AAL31551.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58502 MW; D20034269B926E35 CRC64;

```

```

Query Match 99.1%; Score 2868; DB 12; Length 544;
Best Local Similarity 99.3%; Pred. No. 5.5e-229;
Matches 540; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
QY 3 MASKADPQSDAGSAGQQLVPEVNTADPLPMEPVAGTAVATAGVNMIDPWIIVNFVQ 62
Db 1 MASKADPQSDAGSAGQQLVPEVNTADPLPMEPVAGTAVATAGVNMIDPWIIVNFVQ 60
QY 63 SPOQGEFTTSSSLTIAQATLFPFHVADVRLTLEPIEMPLEDVNRVLYHTNDNQPTMR 122
Db 61 SPOQGEFTTSSSLTIAQATLFPFHVADVRLTLEPIEMPLEDVNRVLYHTNDNQPTMR 120

```

```

QY 123 IIIVCCVPPGFTSSSLTIAQATLFPFHVADVRLTLEPIEMPLEDVNRVLYHTNDNQPTMR 182
Db 121 IIIVCCVPPGFTSSSLTIAQATLFPFHVADVRLTLEPIEMPLEDVNRVLYHTNDNQPTMR 180
QY 183 CMLYPLRTGGGSGNSDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPLQ 242
Db 181 CMLYPLRTGGGSGNSDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPLQ 240
243 SNRSPSSLIQGMILSPDASQVVFQNGRCLIDGQLGTTTATSGQLFRVRGKINQAR 302
241 SNRSPSSLIQGMILSPDASQVVFQNGRCLIDGQLGTTTATSGQLFRVRGKINQAR 300
303 NLTEVDGKPFMAFDSAPVGFDFGKCDWHMRISKTPTNNTSSGDPMRSVSVQTNVQGF 362
301 NLTEVDGKPFMAFDSAPVGFDFGKCDWHMRISKTPTNNTSSGDPMRSVSVQTNVQGF 360
363 HLGSIQFDEVFNHPTGDIYGTIEWISQPTPGTDLINLWEIPDYGSSLSQAANLAPPV 422
361 HLGSIQFDEVFNHPTGDIYGTIEWISQPTPGTDLINLWEIPDYGSSLSQAANLAPPV 420
423 FPGFGEALYVFSAPFPNNRSPNDVPCLLPQEIYTHFVSEQAPTMGDALHLYVDP 482
421 FPGFGEALYVFSAPFPNNRSPNDVPCLLPQEIYTHFVSEQAPTMGDALHLYVDP 480
483 RNLGFEKLYPGGYLTCVPGNGVAGPQQLPLNGVLFVSWVSRYQLKPVGTASTARSG 542
481 RNLGFEKLYPGGYLTCVPGNGVAGPQQLPLNGVLFVSWVSRYQLKPVGTASTARSG 540
543 VRR 546
541 VRR 544

```

RESULT 3

```

Q9IV47 PRELIMINARY; PRT; 546 AA.
AC Q9IV47;
DT 01-OCT-2000 (TRENBLrel. 15, Created)
DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Whitrose/96/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122914;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=HU/NLV/Whitrose/96/UK;
RA MEDLINE=20404883; PubMed=1094950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL; AJ277610; CAB89091.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 546 AA; 58737 MW; 501F9D2BE2A77B51 CRC64;

```

```

Query Match 99.0%; Score 2865; DB 12; Length 546;
Best Local Similarity 98.9%; Pred. No. 9.7e-229;
Matches 540; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 1 MMASKADPQSDAGSAGQQLVPEVNTADPLPMEPVAGTAVATAGVNMIDPWIIVNF 60
Db 1 MMASKADPQSDAGSAGQQLVPEVNTADPLPMEPVAGTAVATAGVNMIDPWIIVNF 60

```

```

QY 61 VQSQGEFTTSSSLTIAQATLFPFHVADVRLTLEPIEMPLEDVNRVLYHTNDNQPTMR 120
Db 61 VQSQGEFTTSSSLTIAQATLFPFHVADVRLTLEPIEMPLEDVNRVLYHTNDNQPTMR 120
121 GKIIIVCCVPPGFTSSSLTIAQATLFPFHVADVRLTLEPIEMPLEDVNRVLYHTNDNQPTMR 180

```



```

Db      121  GKIIVCCVPGFTSSSLTIAQATLFPHVADVRLTLEPIEMPLEDVNRVLYHTNDSTQPTMR 180
QY      181  LVCMLYTLRTGGSGNSDSFVVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPQ 240
Db      181  LVCMLYTLRTGGSGNSDSFVVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPQ 240
QY      241  TILSNRFPPLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGAR 300
Db      241  TILSNRFPPLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGAR 300
QY      301  TILNLTVDGKPFMAFDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPQ 360
Db      301  TILNLTVDGKPFMAFDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPQ 360
QY      361  VPHLSIQDFEVFNHTGDIYGTIEWISQFSTPPGTDINLWEIPDYGSSLSQAANLAPV 420
Db      361  VPHLSIQDFEVFNHTGDIYGTIEWISQFSTPPGTDINLWEIPDYGSSLSQAANLAPV 420
QY      421  FPPGFEALVYFVSAPPGNNRSAPNDVPCLLPQEVYTHFVSEQAPTMDGDAALLHYVDP 480
Db      421  FPPGFEALVYFVSAPPGNNRSAPNDVPCLLPQEVYTHFVSEQAPTMDGDAALLHYVDP 480
QY      481  TNRNLGEFKLYPGGYLTCVPGNGVAGAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTAR 540
Db      481  TNRNLGEFKLYPGGYLTCVPGNGVAGAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTAR 540
QY      541  LGVRR 546
Db      541  LGVRR 546

RESULT 4
Q9DU46
ID Q9DU46 PRELIMINARY; PRT; 544 AA.
AC Q9DU46;
DT 01-MAR-2001 (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=99565;
RN [1] SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Chiba 407/1987/JP;
RA Someya Y., Takeda N., Miyamura T.;
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.
RN [2] SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Chiba 407/1987/JP;
RX MEDLINE=20569531; PubMed=11118371;
RA Someya Y., Takeda N., Miyamura T.;
RT "Complete nucleotide sequence of the Chiba virus genome and functional expression of the 3C-like protease in Escherichia coli.";
RL Virology 278:490-500(2000).
DR EMBL; AB042808; BAB10267.1;
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58351 MW; 1DB7A4E81FBB246F CRC64;

```

```

Query Match 77.1%; Score 2231.5; DB 12; Length 544;
Best Local Similarity 74.7%; Pred. No. 3.1e-176;
Matches 407; Conservative 59; Mismatches 78; Indels 1; Gaps 1;

QY 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWINNF 60
Db 1 MMASKDAPSDAGATGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWINNF 60
QY 61 VOSPOGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMVRVVLGNAFSA 120
Db 61 VOSPOGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMVRVVLGNAFSA 120

```

```

Db      61  VOAPOGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGVGNMVRVVLGNAFSA 120
QY      121  GKIIVCCVPGFTSSSLTIAQATLFPHVADVRLTLEPIEMPLEDVNRVLYHTNDSTQPTMR 180
Db      121  GKIIVCCVPGFTSSSLTIAQATLFPHVADVRLTLEPIEMPLEDVNRVLYHTNDSTQPTMR 180
QY      181  LVCMLYTLRTGGSGNSDSFVVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPQ 240
Db      181  LVCMLYTLRTGGSGNSDSFVVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPQ 240
QY      241  TILSNRFPPLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGAR 300
Db      241  TILSNRFPPLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQGAR 300
QY      301  TILNLTVDGKPFMAFDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPQ 360
Db      301  TILNLTVDGKPFMAFDSFVAGRVLTAPSSDFSLFVLPPTIEQKTRAFVTPNIPQ 360
QY      361  VPHLSIQDFEVFNHTGDIYGTIEWISQFSTPPGTDINLWEIPDYGSSLSQAANLAPV 420
Db      361  VPHLSIQDFEVFNHTGDIYGTIEWISQFSTPPGTDINLWEIPDYGSSLSQAANLAPV 420
QY      421  FPPGFEALVYFVSAPPGNNRSAPNDVPCLLPQEVYTHFVSEQAPTMDGDAALLHYVDP 480
Db      421  FPPGFEALVYFVSAPPGNNRSAPNDVPCLLPQEVYTHFVSEQAPTMDGDAALLHYVDP 480
QY      481  TNRNLGEFKLYPGGYLTCVPGNGVAGAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTAR 540
Db      481  TNRNLGEFKLYPGGYLTCVPGNGVAGAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTAR 540
QY      541  LGVRR 545
Db      541  LGVRR 545

RESULT 5
Q9QTE7
ID Q9QTE7 PRELIMINARY; PRT; 544 AA.
AC Q9QTE7;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Capsid protein.
OS Chiba virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=99565;
RN [1] SEQUENCE FROM N.A.
RC STRAIN=Chiba 407;
RA Natori K., Suzuki K., Yamakawa Y., Tatsumi M., Sakae K., Kobayashi S., Shinozaki K., Ishiko H., Miyamura T., Takeda N.;
RT "Expression and self-assembly of capsid proteins of the Chiba virus, a genetically distinct Norwalk-like virus.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB022679; BAA82106.1;
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58323 MW; ABD1C1FC4F93D872 CRC64;

Query Match 77.0%; Score 2228.5; DB 12; Length 544;
Best Local Similarity 74.5%; Pred. No. 5.5e-176;
Matches 406; Conservative 59; Mismatches 79; Indels 1; Gaps 1;

QY 1 MMASKDAPQSDAGSAGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWINNF 60
Db 1 MMASKDAPSDAGATGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWINNF 60
QY 61 VOSPOGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMVRVVLGNAFSA 120
Db 61 VOSPOGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMVRVVLGNAFSA 120

```

Wed Jun 2 09:13:28 2004

us-09-926-799-3.rspt

121 GKIIIVCCVPPGFTSSSLTIAOATLPHVIAVRTLEPIEMPLEDVRNVLXHTNDNQPTMR 180  
121 GKVIICCVPPGQSRSLTIAOATLPHVIAVRTLDPEVEPLEDVRNVLXHTNDQPTMR 180  
181 LVCMLYPLRTGGSGNSDSFVAVGRVLTAPSSDFSLFPLVPPTEQKTRFTVNIPLQ 240  
181 LLCMLYPLRTGGSGGSDSFVAVGRVLTCPGPDNFLFLVPPTEQKTRFTVNIPLK 240  
241 TILNSRFPSSLIOGMLSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQAR 300  
241 YLSNSRIPNPIEGMSLSPDQTNQVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQAR 300  
301 TILNTEVDGKPFMAFDPAPVFPDFGKCDHMRISKTPNNNTSSGDPMRSVSVQTNVQGF 360  
301 VLNLTELDGSPFMAFDPAPVFPDFGKCDHMRISKTPNNNTSSGDPMRSVSVQTNVQGF 360  
361 VPHLSIQDEVENHHTGYIGTIWISQPSPPGTDINLWIPDYGSSLSQAANLAPPV 420  
361 VPHLSITLDDNVS-SGGDYIGTIQWTSPPDSGGGANTFWKIPDYGSSLSQAANLAPPV 419  
421 FPGFGEALYFVSAPFPGPNRNSAPNDVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDPD 480  
420 YPGFNEVIYFVMAFPGPNQSGSNLVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDPD 479  
481 TNRNLGFEKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARSR 540  
480 TNRNLGFEKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARSR 539  
541 LGVRR 545  
540 LGVRR 544

## RESULT 6

Q91185 PRELIMINARY; PRT; 544 AA.  
AC Q91185;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus HU/NLV/Koblentz/433/2000/DE.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=165252;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HU/NLV/Koblentz 433/2000/DE;  
RA Kuenkel U., Schreier E.;  
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany."  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF394960; AAK72048.1;  
DR InterPro; IPR004005; Calici.coat.  
DR InterPro; IPR008975; Viral.Cap.coat.  
DR Pfam; PF00915; Calici.coat.1;  
SQ SEQUENCE 544 AA; 58222 MW; 1B39A595B733A6A8 CRC64;

Query Match 76.5%; Score 2215.5; DB 12; Length 544;  
Best Local Similarity 74.1%; Pred. No. 6.6e-175;  
Matches 404; Conservative 60; Mismatches 80; Indels 1; Gaps 1;

QY 1 MMASKADPOSADGAGAGQLVPEVNTADPLPMEPVAGFTTAVATAGOVNMDPWINNF 60  
DB 1 MMASKADTPSADGATGAGQLVPEVNTAGTIPIDPVAGSSTALATAGVNLIDPWINNF 60  
QY 61 VSPQCEFTISNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMVRVILLAGNAPSA 120  
DB 61 VQAPQCEFTISNNTPGDVLDLQGLPHLNPFLSHLSQMYNGVGNMVRVILLAGNAPSA 120  
QY 121 GKIIIVCCVPPGFTSSSLTIAOATLPHVIAVRTLEPIEMPLEDVRNVLXHTNDNQPTMR 180  
DB 121 GKVIICCVPPGQSRSLTIAOATLPHVIAVRTLDPEVEPLEDVRNVLXHTNDQPTMR 180

QY 181 LVCMLYPLRTGGSGNSDSFVAVGRVLTAPSSDFSLFPLVPPTEQKTRFTVNIPLQ 240  
DB 181 LLCMLYPLRTGGSGGSDSFVAVGRVLTCPGPDNFLFLVPPTEQKTRFTVNIPLK 240  
QY 241 TILNSRFPSSLIOGMLSPDASQVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQAR 300  
DB 241 YLSNSRIPNPIEGMSLSPDQTNQVQFQNGRCLIDGQLLGTTPATSGQLFRVRGKINQAR 300  
QY 301 TILNTEVDGKPFMAFDPAPVFPDFGKCDHMRISKTPNNNTSSGDPMRSVSVQTNVQGF 360  
DB 301 VLNLTELDGSPFMAFDPAPVFPDFGKCDHMRISKTPNNNTSSGDPMRSVSVQTNVQGF 360  
QY 361 VPHLSIQDEVENHHTGYIGTIWISQPSPPGTDINLWIPDYGSSLSQAANLAPPV 420  
DB 361 VPHLSITLDDNVS-SGGDYIGTIQWTSPPDSGGGANTFWKIPDYGSSLSQAANLAPPV 419  
QY 421 FPGFGEALYFVSAPFPGPNRNSAPNDVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDPD 480  
DB 420 YPGFNEVIYFVMAFPGPNQSGSNLVPCLLPQEIYTHFVSEQAPTMGDAALLHYVDPD 479  
QY 481 TNRNLGFEKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARSR 540  
DB 480 TNRNLGFEKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWSRFYQLKPVGTASTARSR 539  
QY 541 LGVRR 545  
DB 540 LGVRR 544

## RESULT 7

Q918C5 PRELIMINARY; PRT; 544 AA.  
AC Q918C5;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
GN ORF2.  
OS Norwalk-like virus NLV/Baltimore/277/1993/US.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=171836;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HU/NLV/Baltimore/277/1993/US;  
RX MEDLINE=97193806; PubMed=9041391;  
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
RT "A one-tube method of reverse transcription-PCR to efficiently amplify a 3-kilobase region from the RNA polymerase gene to the poly(A) tail of small round-structured viruses (Norwalk-like viruses).";  
RL J. Clin. Microbiol. 35:570-577(1997).  
RN [2]  
RP SEQUENCE OF 100-192 FROM N.A.  
RC STRAIN=HU/NLV/Baltimore/277/1993/US;  
RX MEDLINE=98071277; PubMed=9407386;  
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K., Seto Y., Monroe S.S., Glass R.I.;  
RT "Correlation of patient immune responses with genetically characterized small round-structured viruses involved in outbreaks of nonbacterial acute gastroenteritis in the United States, 1990 to 1995.";  
RL J. Med. Virol. 53:372-383(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HU/NLV/Baltimore/277/1993/US;  
RX MEDLINE=20266071; PubMed=10804147;  
RA Ando T., Noel J.S., Fankhauser R.L.;  
RT "Genetic classification of 'Norwalk'-like viruses.";  
RL J. Infect. Dis. 181:S336-S348(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HU/NLV/Baltimore/277/1993/US;

RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF414404; AAL12865.1; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR Pfam; PF00915; Calici\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 544 AA; 58208 MW; D5170FA4B958A672 CRC64;

Query Match 76.4%; Score 2210.5; DB 12; Length 544;  
Best Local Similarity 73.9%; Pred. No. 1.7e-174;  
Matches 403; Conservative 61; Mismatches 80; Indels 1; Gaps 1;

QY 1 MMASKADPQSDAGSAGAGOLVPEVNTADPLMEPVAGPTTAVATAGVNMIDPVIWNP 60  
DB 1 MMASKADTPSDAGAGAGOLVPEVNTADPIPDPVAGSTALATAGVNLIDPVIWNP 60  
QY 61 VQSPQGEFTISPNNTPGDIPLDQLGPHLNPFLSHLSQMYNGWGNMVRVILLAGNAFSA 120  
DB 61 VQAPQGEFTISPNNTPGDIPLDQLGPHLNPFLSHLSQMYNGWGNMVRVILLAGNAFSA 120  
QY 121 GKIIICVCPGPGTSSSLTIAQATLFPHVIAVRLTPEIEMPLEDVNRVLYHTNDNQPTWR 180  
DB 121 GKVIIICVCPGPGFSRTLSIAQATLFPHVIAVRLTPEIEMPLEDVNRVLYHTNDNQPTWR 180  
QY 181 LVCMLYTPLRTGGSGNSDSFVAGRVLTAPSDFSFLVPPTEIOKTRAFVVPNIPLQ 240  
DB 181 LLCMLYTPLRTGGSGNSDSFVSGRVLTAPSDFSFLVPPTEIOKTRAFVVPNIPLK 240  
QY 241 TLSNSRFPSSLIOQMILSPDASQVQFQNGRCLIDGLGTTTATSGQLFRVRGKINQGAR 300  
DB 241 CLNSRIPIPIEGMSLSPOTQNVQFQNGRCLIDGLGTTTATSGQLFRVRGKINQGAR 300  
QY 301 TLMLETVGDKPMAPFSPAPVPGPFQKCDWMHRIKSTPNNTSSGDPMSRSVQTNVQGF 360  
DB 301 VLNLTELDGSPFMAFAAPAGPDLGSCDWHIEMSKIPNSSQNNPIVNSVKPNSQGF 360  
QY 361 VPHLSGIQFDEVFNHTDYGITIEWISQPS-TPGTDINLWEIPDYGSSLSQANLAPPV 420  
DB 361 VPHLSGITLDENVIS-SCGDYIGITQWTSPPSDGGANTFWKIPDYGSSLSQALAPAV 419  
QY 421 FPPGFGALVYFVSAPFPGNNRSPNDVPCLLPOEYITHFVSEQATMGDAALLHYVDP 480  
DB 420 YPPGFNEVIVFASIPGNSQSPNLVPCLLPOEYITHFVSEQATMGDAALLHYVDP 479  
QY 481 TNRNLGEFKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWRSRFLQKPVGTASTARS 540  
DB 480 TNRNLGEFKLYPGGYLTCVPGNSSTGQQLPLDGVFVFASWRSRFLQKPVGTAGPGR 539  
QY 541 LGVRR 545  
DB 540 LGVRR 544

RESULT 8  
Q91V43 ID Q91V43 PRELIMINARY; PRT; 543 AA.  
AC Q91V43;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus HU/NLV/Musgrove/89/UK.  
OC Viruses; serNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=122918;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN=HU/NLV/Musgrove/89/UK;  
RC MEDLINE=20404883; PubMed=10949950;  
RX Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,  
RA Clegg J.C., Chamberlain J., Brown D.W.G.;  
RT 'Capsid protein diversity among 'Norwalk-like' viruses.';  
RL Virus Genes 20:227-236(2000).

DR EMBL; AJ277614; CAB89095.1; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 543 AA; 58216 MW; 11F97A3989CCCS40 CRC64;

Query Match 76.3%; Score 2208; DB 12; Length 543;  
Best Local Similarity 74.0%; Pred. No. 2.7e-174;  
Matches 404; Conservative 63; Mismatches 75; Indels 4; Gaps 3;

QY 1 MMASKADPQSDAGSAGAGOLVPEVNTADPLMEPVAGPTTAVATAGVNMIDPVIWNP 60  
DB 1 MMASKADTPSDAGAGAGOLVPEVNTADPIPDPVAGSTALATAGVNMIDPVIWNP 60  
QY 61 VQSPQGEFTISPNNTPGDIPLDQLGPHLNPFLSHLSQMYNGWGNMVRVILLAGNAFSA 120  
DB 61 VQAPQGEFTISPNNTPGDIPLDQLGPHLNPFLSHLSQMYNGWGNMVRVILLAGNAFSA 120  
QY 121 GKIIICVCPGPGTSSSLTIAQATLFPHVIAVRLTPEIEMPLEDVNRVLYHTNDNQPTWR 180  
DB 121 GKVIIICVCPGPGFSRTLSIAQATLFPHVIAVRLTPEIEMPLEDVNRVLYHTNDNQPTWR 180  
QY 181 LVCMLYTPLRTGGSGNSDSFVAGRVLTAPSDFSFLVPPTEIOKTRAFVVPNIPLQ 240  
DB 181 LLCMLYTPLRTGGSGNSDSFVAGRVLTAPSDFSFLVPPTEIOKTRAFVVPNIPLQ 240  
QY 241 TLSNSRFPSSLIOQMILSPDASQVQFQNGRCLIDGLGTTTATSGQLFRVRGKINQGAR 300  
DB 241 LLSNSRVNLIQMSVLSPDQANVQFQNGRCLIDGLGTTTATSGQLFRVRGKINQGAR 300  
QY 301 TLMLETVGDKPMAPFSPAPVPGPFQKCDWMHRIKSTPNNTSSGDPMSRSVQTNVQGF 360  
DB 301 VINLTELDSPPFAEAPAPGPDGTSMDWHIEMSLNSQSSGNPILRLDIQPNSSDF 360  
QY 361 VPHLSGIQFDEVFNHTDYGITIEWISQPS-TPGTDINLWEIPDYGSSLSQANLAPP 419  
DB 361 VPHLSGSVTTAID-TAGDTLGTIQTSPNSVTPVDVNFVTIIPDYGSSLSQALAPV 419  
QY 420 FPPGFGALVYFVSAPFPGNNRSPNDVPCLLPOEYITHFVSEQATMGDAALLHYVDP 479  
DB 420 YPPGFNEVIVFASIPGNSQSPNLVPCLLPOEYITHFVSEQATMGDAALLHYVDP 479  
QY 480 TNRNLGEFKLYPGGYLTCVPGNGVAGPQQLPLNGVFLFVSWRSRFLQKPVGTASTARS 539  
DB 480 TNRNLGEFKLYPGEFTICVFN-CTGPOQLPLNGVVFASWRSRFLQKPVGTASSARG 537  
QY 540 LGVRR 545  
DB 538 LGVRR 543

RESULT 9  
Q91EJ2 ID Q91EJ2 PRELIMINARY; PRT; 544 AA.  
AC Q91EJ2;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Nucleocapsid protein.  
OS Human calicivirus HU/NLV/Queen's Arms/Leeds/92/UK.  
OC Viruses; serNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=159480;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Clegg C.S., Chamberlain J., Green J., Brown D.W.G., Lewis D.;  
RT "Analysis of diversity in the capsid and ORF3 proteins of some  
RT Norwalk-like viruses circulating in the U.K.";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AJ313030; CAC40987.1; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.

RA	Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K., Seto Y., Monroe S.S., Glass R.I.; "Correlation of patient immune responses with genetically characterized small round-structured viruses involved in outbreaks of nonbacterial acute gastroenteritis in the United States, 1990 to 1995.";
RT	J. Med. Virol. 53:372-383(1997).
RL	[3]
RN	SEQUENCE FROM N.A.
RP	STRAIN=Hu/NL/V/Appalachicola Bay/318/1995/US;
RC	MEDLINE=20266071; PubMed=10804147;
RX	Ando T., Seto Y., Noel J.S., Fankhauser R.L.;
RA	EMBL; AF141406; AAL12971.1; -
RR	InterPro; IPR004005; Calici.coat.
DR	InterPro; IPR008975; Viral.Cap.coat.
DR	Pfam; PF00915; Calici.coat.1.
DR	SEQUENCE 543 AA; 58362 MW; 2C0A035E80A20EF5 CRC64;
SQ	
Query Match	76.1%; Score 2202; DB 12; Length 543;
Best Local Similarity	74.0%; Pred. No. 8.6e-174;
Matches	404; Conservative 63; Mismatches 75; Indels 4; Gaps 3;
QY	1 MMASKADAPQSADGASGAGQLVPEVTADPLPMBPVAGPTTAVATAGOVNMIDPWVNNF 60
DB	1 MMASKADSPSADGANGAGQLVPEVNNAEPLDLPVAGASTALATAGOVNMIDPWVNNF 60
QY	61 VOSPOGEFTISPNNTPGDILFDLQGLPHNLFPFLSHLSOMYNGWGNMVRILLAGNAFSA 120
DB	61 VQAPOGEFTISPNNTPGDILFDLQGLPHNLFPFLSHLSOMYNGWGNMVRILLAGNAFTA 120
QY	121 GKIIICCVPPGTSSSLTAQTLPFHVIADVRLTEPIEMPLEDVNRVLYHTNDNQPTMR 180
DB	121 GKIVICCVPPGQSRSLTAQTLPFHVIADVRLTEPIEMPLEDVNRVLYHTNDNQPTMR 180
QY	181 LVCMLYTLRTPRTGGSGNSDSFVVAGRVLTPASSDFSFLVPPTIEOKTRAFVNPILQ 240
DB	181 LLMCLYTPLRTLRTGGSGGTDFAVVAGRVLTCPSDDNFLLFVPPVEQKTRPFVSVPNIPLQ 240
QY	241 TLNSRFPSSLIQGMILSPDASQVQFONGRCLIDQLLGTTTPATSCQLFRVRGKINOGAR 300
DB	241 LLNSRVENLIQSMVSPDOAQNVQFONGRCTDQLLGTTFVSVSQLIKFRGVKSAGSR 300
QY	301 TMLNLTVDGKPFMAFSDPAPVGPPDFGKCDWMRIKTPNTTSSGDPMRSVSVQTNVQGF 360
DB	301 VINLTLDGSPFLAEAPAPTGPDLGTSMDHVMESLNSNSSQSNPILLRDIQNSSDF 360
QY	361 VPHLGSIOFDEVNHPTGYDTIETIEWISOPS--TPCTDINLWEIPDYGSSLSQAANLAPP 419
DB	361 VPHLGSVAVTTAID-VAGDYTGTIQMTSQPSNVTPVPDVFNTFIPOYGSLAEASQLAPV 419
QY	420 VFPPGFEALVYFVSAPFGPNRNASPDVPCLLPQEYITHFVSEQAPTMGDAALLHYDP 479
DB	420 VYPGFEALVYFVMSPIPGENTAHKNLVPRLLPQEFTVTHFVSEQAPSMEALVHYDP 479
QY	480 DTNRNLGEFKLYPGGYLTCVPNGVAGAPOLPLNGVFLFVSWVSRFYQLKPVGTAstars 539
DB	480 DTNRNLGEFKLYPEGFTCMEN--GTGPQOOLPLNGVVFVFWASWSRFYQLKPVGTAstars 537
QY	540 RLGVRR 545
DB	538 RLGVRR 543
RESULT 11	
ID	Q918D1 PRELIMINARY; PRT; 544 AA.

```
AC Q918D1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
GN Norwalk-like virus NLV/New Orleans/266/1993/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171834;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/266/1993/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/266/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of Norwalk-like viruses.";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/266/1993/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414402; AAL12959.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat_1.
SQ SEQUENCE 544 AA; 58354 MW; 50387E99852E801D CRC64;

Query Match 76.0%; Score 2201.5; DB 12; Length 544;
Best Local Similarity 73.9%; Pred. No. 9.5e-174;
Matches 403; Conservative 58; Mismatches 83; Indels 1; Gaps 1;

QY 1 MMASKDAPQSGADGAGAGQVPEVNTADPLPMEPVAGTTAVATAGVNMIDPVIWNNF 60
DB 1 MMASKDAPQSGADGAGAGQVPEVNTADPLPMEPVAGTTAVATAGVNMIDPVIWNNF 60
QY 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMVRVILLAGNAFSA 120
DB 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMVRVILLAGNAFSA 120
QY 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMVRVILLAGNAFSA 120
DB 61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMVRVILLAGNAFSA 120
QY 121 GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNVLHYTNDNQPTMR 180
DB 121 GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLTLEPIEMPLEDVRNVLHYTNDNQPTMR 180
QY 181 LVCMLYTPLRTGGSGNSDSFVAVGRVLTAPSSDFPLVPPTIEQKTRAFVTPNIPLO 240
DB 181 LVCMLYTPLRTGGSGNSDSFVAVGRVLTAPSSDFPLVPPTIEQKTRAFVTPNIPLO 240
QY 181 LLCMLYTPLRTGGSGNSDSFVAVGRVLTAPSSDFPLVPPTIEQKTRAFVTPNIPLO 240
DB 181 LLCMLYTPLRTGGSGNSDSFVAVGRVLTAPSSDFPLVPPTIEQKTRAFVTPNIPLO 240
QY 241 TILSNRFPESLIQGMILSPDASQVQFQNGRCILDGQLLGTTPATSGQLFRVRGKINQCAR 300
DB 241 TILSNRFPESLIQGMILSPDASQVQFQNGRCILDGQLLGTTPATSGQLFRVRGKINQCAR 300
QY 301 TLNLTVDGKPFMAFDPSPAPVGFPPDFGKCDWHMIRISKTPNNTSSGDPNRSVSQTNVQGF 360
DB 301 TLNLTVDGKPFMAFDPSPAPVGFPPDFGKCDWHMIRISKTPNNTSSGDPNRSVSQTNVQGF 360
QY 301 ALNLTLDGSPFMAFDPSPAPVGFPPDFGKCDWHMIRISKTPNNTSSGDPNRSVSQTNVQGF 360
DB 301 ALNLTLDGSPFMAFDPSPAPVGFPPDFGKCDWHMIRISKTPNNTSSGDPNRSVSQTNVQGF 360
QY 361 VPHLGSIQFDE-VFNHPTGDIYGTIEW-----ISQSTPPCTDNLNWEIPDYGSSLSQA 413
DB 361 VPHLGSITLDENVS--GGDYIGNIQWTLILLVILEPKT-----NFKIPDYGSSLSQA 413
QY 414 ANLAPVPPFPGFEALVYFVSAPFGPNRNSAPNDVPCLLPQEIYTHFVSEQAPTMDAAL 473
DB 414 ANLAPVPPFPGFEALVYFVSAPFGPNRNSAPNDVPCLLPQEIYTHFVSEQAPTMDAAL 473
QY 413 SOLAPVPPFPGFEALVYFVSAPFGPNRNSAPNDVPCLLPQEIYTHFVSEQAPIGEAAL 472
DB 413 SOLAPVPPFPGFEALVYFVSAPFGPNRNSAPNDVPCLLPQEIYTHFVSEQAPIGEAAL 472
QY 474 LHYVDPDTRNLGEPKLYPGGYLTCPVNGVAGGPOQLPLNGVFLPVSWSRYPQLKPVGT 533
DB 474 LHYVDPDTRNLGEPKLYPGGYLTCPVNGVAGGPOQLPLNGVFLPVSWSRYPQLKPVGT 533
```

Qy	541	LGVR	545	
Db	540	LGVR	544	
RESULT 14				
ID	Q8JXJ0	PRELIMINARY;	PRT;	543 AA.
AC	Q8JXJ0;			
DT	01-OCT-2002 (Tremblrel. 22, Created)			
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)			
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)			
DE	Capsid.			
OS	Norwalk-like virus.			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;			
OC	Norovirus.			
OX	NCBI_TaxID=95340;			
RN	[1] _			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SzUG1;			
RA	Katayama K., Kojima S.;			
RL	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.			
RN	[2] _			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=SzUG1;			
RX	MEDLINE=22192455; PubMed=1202225;			
RA	Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,			
RA	Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,			
RA	Gojobori T., Takeda N.;			
RT	"Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like			
RT	viruses.";			
RL	Virolgy 299:225-239 (2002).			
DR	EMBL; AB039774; BAC11813.1; -.			
DR	InterPro: IPR004005; Calici coat.			
DR	InterPro: IPR008975; Viral Cap_coat.			
DR	Pfam; PF00915; Calici coat; 1.			
SQ	SEQUENCE 543 AA; 58761 MW; E0DE7489A9C3488C CRC64;			
Query Match 75.2%; Score 2178; DB 12; Length 543;				
Best Local Similarity 73.3%; Pred. No. 8.4e-172;				
Matches 401; Conservative 65; Mismatches 75; Indels 6; Gaps 4;				
Qy	1	MMASKDAPQADGASGAGQLVPEVNTADPLPEPVAGPTTAVATAGVNMIDPWVNNF	60	
Db	1	MMASKDAPQADGANGAGQLVPEVNTADPLPEPVAGPTTAVATAGVNMIDPWVNNF	60	
Qy	61	VSPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWGNMVRVILAGNAFSA	120	
Db	61	VQAPQGEFTISPNNTPGDILFDLQGLPHLPFLSHLSQMYNGWGNMVRVILAGNAFTA	120	
Qy	121	GKIIIVCCVPPGFTSSLTIAQATLFPVHVIADVRLTLEPIEMPLEDVRNLYHTNDNQPTMR	180	
Db	121	GKVIICCVPPGFTSSLTIAQATLFPVHVIADVRLTLEPIEMPLEDVRNLYHTNDNQPTMR	180	
Qy	181	LVCMLYTLRTGGSGNSDSFVAVGRVLTAPSSDFSLFVLPPTIEQKTRAFVVPNIPLQ	240	
Db	181	LLCMLYTLRTGGSGSGTDAFVAVGRVLTCPSPDFNFLFLVPTVEQKTRFSPVNIPLQ	240	
Qy	241	TLNSRPPSLIQCMILSPDASQVQFQNGRCLIDQLLGTTPATSGQLFRVRGKINOGAR	300	
Db	241	NLSNRVPSLIQSMVLSNDHQAQVQFQNGRCLTDGHLGLTTPVSAGQLMKFRKVTSGSK	300	
Qy	301	TLNLTVDGKPFMAFDSAPVGFDFGKCDHMRISKTPNNTSSGDMRSVSQTNVQGF	360	
Db	301	VLNLTVDGSPFLFAEPAPAGPFDLGKCDHVMESLYQDN-NHAPDPIVLHAIENPSSSF	359	
Qy	361	VPHLSIQDFEVNHTGDIYGTIEWISQSTP--FCTDINLWEIPDYGSSLSQAANLAP	418	
Db	360	VPHLSVSFKENVN-VAGDYVCTIQTSPSPDPHPNADVDWFSPIDYGSNLAESAQLAP	418	
Qy	419	PVPPGFGEALVYFVSAPFGPNRNSAPNDVPCLLPQEIYTHFVSEQAPTMGDAALLHYVD	478	
Db	419	VVYPPGFGEALVYFVSAPFGPNRNLNLPVCPQOEFGHFVSEQAPALGEAALLHYVD	478	



**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 46.7329 Seconds  
(without alignments)  
3289.030 Million cell updates/sec

Title: US-09-926-799-4  
Perfect score: 2891  
Sequence: 1 MMASKATPDSADGATGAGQ.....YOLKPVGTAGPARGRLGVRR 544

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2891	100.0	544	4	AAB49703 Small rou
2	2231.5	77.2	546	4	AAB49702 Small rou
3	2100	72.6	530	2	AAR50972 Norwalk v
4	2100	72.6	530	7	Adc72176 Norwalk v
5	2093	72.4	530	4	AAB49701 Small rou
6	2054	71.0	530	2	AAR57091 Small rou
7	1979	68.5	545	4	AAB49700 Small rou
8	1170.5	40.5	540	4	AAB49706 Small rou
9	1165.5	40.3	542	4	AAB49708 Small rou
10	1151.5	39.8	539	4	AAB49704 Small rou
11	1145	39.6	535	4	AAB49707 Small rou
12	1143.5	39.6	548	4	AAB49705 Small rou
13	1134	39.2	548	5	AAU91272 Norwalk v
14	1124.5	38.9	550	4	AAB49709 Small rou
15	970	33.6	541	4	AAB49710 Small rou
16	309	10.7	579	2	AAW08143 RHDV caps
17	268	9.3	568	4	AAB67462 Amino aci
18	264.5	9.1	547	4	AAW50108 Feline ca
19	264.5	9.1	671	4	AAW50107 Feline ca
20	263.5	9.1	668	2	AAR10686 Feline ca
21	263.5	9.1	668	4	AAE04304 Feline ca
22	263.5	9.1	669	4	AAB67461 Amino aci
23	261.5	9.0	623	4	AAB47044 Feline Ca
24	261.5	9.0	623	4	AAB47043 Feline Ca
25	250.5	8.7	622	4	AAB47045 Feline Ca

26	143.5	5.0	2971	7	ADC35084	Adc35084 Human bre
27	143.5	5.0	2972	4	AAB50363	Aab50363 Human SRC
28	143.5	5.0	3118	4	AAB50362	Aab50362 Human SRC
29	142.5	4.9	2735	6	ABR41356	ABr41356 Human DIT
30	141	4.9	2206	2	AAR22210	Aar22210 True type
31	140.5	4.9	2971	3	AAB41231	Aab41231 Human ORF
32	140	4.8	40	5	AAU91273	AAu91273 Norwalk v
33	130	4.5	3930	6	ABU18893	ABu18893 Protein e
34	127	4.4	1127	4	AAB95541	Aab95541 Human pro
35	126.5	4.4	1185	6	ABU18969	ABu18969 Protein e
36	126	4.4	40	5	AAU91274	AAu91274 Norwalk v
37	125.5	4.3	2358	6	ABU17839	ABu17839 Protein e
38	125	4.3	1152	7	ADE59446	AdE59446 Human Pro
39	123.5	4.3	6310	6	ABU39869	ABu39869 Protein e
40	121	4.2	1150	4	AAW40399	AAw40399 Human pol
41	121	4.2	2183	3	AAW03533	AAw03533 Murine fa
42	120.5	4.2	895	3	AAV56804	AAv56804 Clathrin
43	117.5	4.1	1155	2	AAR28047	Aar28047 IRS-1 pro
44	117.5	4.1	2016	4	ABW63911	ABw63911 Drosophill
45	117	4.0	1564	7	ADC86801	Adc86801 Human GPC

## ALIGNMENTS

RESULT 1  
AAB49703  
ID AAB49703 standard; protein; 544 AA.  
XX  
AC AAB49703;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 4.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX  
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
(DENK-) DENKA SEIKEN KK.  
XX  
Takeda N, Natōri K, Miyamura T, Kamata K, Sato T, Sato S;  
WPI; 2001-080848/09.  
DR N-PSDB; AAF29144.  
XX  
Kit for the detection and typing of small round-structured virus (SRSV)  
strains for investigation of food poisoning outbreaks, contains  
antibodies.  
XX  
Claim 1; Page 47-49; 84pp; Japanese.  
XX  
This invention relates to a kit for the detection and typing of small  
round structured virus (SRSV) strains. The kit contains antibodies  
directed against peptides represented in sequences AAB49700 - AAB49710,  
which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
used for detecting and typing strains of SRSV in order to prevent the  
spread of infection and to examine the epidemiology of outbreaks  
XX  
Sequence 544 AA;  
Query Match 100.0%; Score 2891; DB 4; Length 544;  
Best Local Similarity 100.0%; Pred. No. 9.4e-249;

Wed Jun 2 09:13:28 2004

Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMASKDATPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIIINNF 60  
DB 1 MMASKDATPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIIINNF 60

QY 61 VQAPQGEFTISPNPTGDLQGLPHLNPFLSHLSQMYNGWGNMRVRVVLAGNAFTA 120  
DB 61 VQAPQGEFTISPNPTGDLQGLPHLNPFLSHLSQMYNGWGNMRVRVVLAGNAFTA 120

QY 121 GKVIICCVPPGQSRITLSIAQATLFPFHVIAVRTLDPEVPLEDRVNLVYHNDTQPTMR 180  
DB 121 GKVIICCVPPGQSRITLSIAQATLFPFHVIAVRTLDPEVPLEDRVNLVYHNDTQPTMR 180

QY 181 LLCMLYTLRTGGASGSDTSFVAVGRVLTCPGDFNFELVPPTEQKTRPFTVPNIPLK 240  
DB 181 LLCMLYTLRTGGASGSDTSFVAVGRVLTCPGDFNFELVPPTEQKTRPFTVPNIPLK 240

QY 241 YLSNSRIPNPIEGMSLSPTQNTQVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSQOR 300  
DB 241 YLSNSRIPNPIEGMSLSPTQNTQVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSQOR 300

QY 301 VLNLTELDSGSPFMAFAAPAPAGPDPDLGSCDWHIEMSKI PNSTQNNPIVTVSVKPNSSQOF 360  
DB 301 VLNLTELDSGSPFMAFAAPAPAGPDPDLGSCDWHIEMSKI PNSTQNNPIVTVSVKPNSSQOF 360

QY 361 VPHLSSTILDENVS-SGGDYIGTIQWTSPPSDSGANTFWKIPDYGSSLAESAQLAPAVY 420  
DB 361 VPHLSSTILDENVS-SGGDYIGTIQWTSPPSDSGANTFWKIPDYGSSLAESAQLAPAVY 420

QY 421 PPGNEVIVVFMAISIPGPNQSGSNLVPCLLPOEYITHFISEQAPTOGEAALLHYVDPT 480  
DB 421 PPGNEVIVVFMAISIPGPNQSGSNLVPCLLPOEYITHFISEQAPTOGEAALLHYVDPT 480

QY 481 NRNLGEPKLYPGGYLTCVPNSSSTGPOQLDGVFVFSWVSRYQLKPVGTAGPARGRL 540  
DB 481 NRNLGEPKLYPGGYLTCVPNSSSTGPOQLDGVFVFSWVSRYQLKPVGTAGPARGRL 540

QY 541 GVR 544  
DB 541 GVR 544

RESULT 2  
ID AAB49702 standard; protein; 546 AA.  
AC AAB49702;  
DT 04-APR-2001 (first entry)  
DE Small round structured virus protein SEQ ID 3.  
XX Small round structured virus; SRSV; food poisoning.  
XX Small round structured virus.  
XX WO200079280-A1.  
XX 28-DEC-2000.  
XX 22-JUN-2000; 2000WO-JP004095.  
XX 22-JUN-1999; 99JP-00175928.  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX (DENK-) DENKA SEIKEN KK.  
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX WPI; 2001-080848/09.  
XX N-PSDB; AAF29143.  
XX

Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

Claim 1; Page 45-47; 84pp; Japanese.

This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

Sequence 546 AA;

Query Match 77.2%; Score 2231.5; DB 4; Length 546;  
Best Local Similarity 74.7%; Pred. No. 6.8e-190;  
Matches 407; Conservative 59; Mismatches 78; Indels 1; Gaps 1;

QY 1 MMASKDATPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIIINNF 60  
DB 1 MMASKDATPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIIINNF 60

QY 61 VQAPQGEFTISPNPTGDLQGLPHLNPFLSHLSQMYNGWGNMRVRVVLAGNAFTA 120  
DB 61 VQAPQGEFTISPNPTGDLQGLPHLNPFLSHLSQMYNGWGNMRVRVVLAGNAFTA 120

QY 121 GKVIICCVPPGQSRITLSIAQATLFPFHVIAVRTLDPEVPLEDRVNLVYHNDTQPTMR 180  
DB 121 GKVIICCVPPGQSRITLSIAQATLFPFHVIAVRTLDPEVPLEDRVNLVYHNDTQPTMR 180

QY 181 LLCMLYTLRTGGASGSDTSFVAVGRVLTCPGDFNFELVPPTEQKTRPFTVPNIPLK 240  
DB 181 LLCMLYTLRTGGASGSDTSFVAVGRVLTCPGDFNFELVPPTEQKTRPFTVPNIPLK 240

QY 241 YLSNSRIPNPIEGMSLSPTQNTQVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSQOR 300  
DB 241 YLSNSRIPNPIEGMSLSPTQNTQVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSQOR 300

QY 301 VLNLTELDSGSPFMAFAAPAPAGPDPDLGSCDWHIEMSKI PNSTQNNPIVTVSVKPNSSQOF 360  
DB 301 VLNLTELDSGSPFMAFAAPAPAGPDPDLGSCDWHIEMSKI PNSTQNNPIVTVSVKPNSSQOF 360

QY 361 VPHLSSTILDENVS-SGGDYIGTIQWTSPPSDSGANTFWKIPDYGSSLAESAQLAPAV 419  
DB 361 VPHLSSTILDENVS-SGGDYIGTIQWTSPPSDSGANTFWKIPDYGSSLAESAQLAPAV 420

QY 420 YPPGNEVIVVFMAISIPGPNQSGSNLVPCLLPOEYITHFISEQAPTOGEAALLHYVDPT 479  
DB 420 YPPGNEVIVVFMAISIPGPNQSGSNLVPCLLPOEYITHFISEQAPTOGEAALLHYVDPT 480

QY 480 TNRNLGEPKLYPGGYLTCVPNSSSTGPOQLDGVFVFSWVSRYQLKPVGTAGPARGR 539  
DB 480 TNRNLGEPKLYPGGYLTCVPNSSSTGPOQLDGVFVFSWVSRYQLKPVGTAGPARGR 540

QY 540 LGVRR 544  
DB 540 LGVRR 544

RESULT 3  
ID AAR50972 standard; protein; 530 AA.  
AC AAR50972;  
DT 16-OCT-2003 (revised)  
DT 25-MAR-2003 (revised)  
DT 05-OCT-1994 (first entry)  
XX Norwalk virus strain 8FIIa protein (encoded by ORF2).  
XX



180 RLLCMLYPLRTGASGSDTSFVAGRVLTCPGPDENFLVLPPTVEQKTRPFTVPNIPL 239  
 181 RLVCMYPLRTGGGTG--DSFVAGRVMTCPSPDENFLVLPPTVEQKTRPFTLPNPL 238  
 240 KYLSNRIPNPIEGMSLSDPOTQNVQFQNGRGTIDGQPLGTPVSVSOLCKFRGRTSQ 299  
 239 SLSLSRAPLPISNGISGPDNVQSVQFQNGRGTIDGRLVGTTPVSLSHVAKIRG--TSNG 296  
 300 RVNLNLTDLGSPFMAFAAPAGFPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQ 359  
 297 TVINLTDLGTPFHPFEGPAPIGFPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQ 351  
 360 FVPHLSITLDENVSSGGYICTIOWTSPDSGGANTFVKIPDYGSSLAESQAPAV 419  
 352 FVPHLSI--QANGIGSGNYIGVLSWSPSPHSGQVDLWKIPNYGSSITEATHLAPSV 409  
 420 YPGFNEVIVFMASIPGPNQSGSNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDP 479  
 410 YPGFGEVLVFFMSKMPG--GAYNL-PCLLPQEIYSHLASEQAPTVEAALLHYVDP 465  
 480 TNRNLGEFKLYPGGYLTCVPNSSTGQOOLPLDGVFVFPASWVSFRFYQLKPVGTAGPARGR 539  
 466 TGRNLGEFKAYPDGFLTCVPNGASSGQOOLPINGVFVFSWVSFRFYQLKPVGTASSARGR 525  
 540 LGVRR 544  
 526 LGLRR 530

RESULT 5  
 AAB49701  
 ID AAB49701 standard; protein; 530 AA.  
 XX  
 AC AAB49701;  
 DT 04-APR-2001 (first entry)  
 XX  
 DE Small round structured virus protein SEQ ID 2.  
 XX  
 DE Small round structured virus; SRSV; food poisoning.  
 XX  
 KW Small round structured virus.  
 XX  
 OS Small round structured virus.  
 XX  
 PN WO200079280-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 22-JUN-2000; 2000WO-JP004095.  
 XX  
 PR 22-JUN-1999; 99JP-00175928.  
 XX  
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 PA (DENK-) DENKA SEIKEN KK.  
 XX  
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 XX  
 DR WPI; 2001-080848/09.  
 DR N-PSDB; AAF29142.  
 XX  
 XX Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.  
 XX  
 PS Claim 1; Page 42-45; 84pp; Japanese.  
 XX  
 XX This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAB20141 -  
 CC AAB20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 CC

Query Match 72.4%; Score 2093; DB 4; Length 530;  
 Best Local Similarity 72.7%; Pred. No. 1.5e-177; Indels 16; Gaps 8;  
 Matches 396; Conservative 53; Mismatches 80;  
 1 MMASKDATPSADGATGAGLVPEVNTADPIPDPVAGSSTALATAGOVNLIIDWIINNF 60  
 1 MMASKDATPSVVDGASGAGLVPEVNASDPLAMPVAGSSTAVATAGOVNPIIDWIINNF 60  
 61 VQAPQGEFTISPNPTGPDVLDLQGLPHNLPFLSHLSOMYNGVMRVRVVLAGNAFTA 120  
 61 VQAPQGEFTISPNPTGPDVLDLQGLPHNLPFLSHLSOMYNGVMRVRVVLAGNAFTA 120  
 121 GKVIICVPPGQSRRTLSIAQATLPFHVITADVRLDPEVPLEDVNRYLNND-TQPTM 179  
 121 GKVIICVPPGQSRRTLSIAQATLPFHVITADVRLDPEVPLEDVNRYLNND-TQPTM 180  
 180 RLLCMLYPLRTGGASGSDTSFVAGRVLTCPGPDENFLVLPPTVEQKTRPFTVPNIPL 239  
 181 RLVCMYPLRTGGGTG--DSFVAGRVMTCPSPDENFLVLPPTVEQKTRPFTLPNPL 238  
 240 KYLSNRIPNPIEGMSLSDPOTQNVQFQNGRGTIDGQPLGTPVSVSOLCKFRGRTSQ 299  
 239 SLSLSRAPLPISNGISGPDNVQSVQFQNGRGTIDGRLVGTTPVSLSHVAKIRG--TSNG 296  
 300 RVNLNLTDLGSPFMAFAAPAGFPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQ 359  
 297 TVINLTDLGTPFHPFEGPAPIGFPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQ 351  
 360 FVPHLSITLDENVSSGGYICTIOWTSPDSGGANTFVKIPDYGSSLAESQAPAV 419  
 352 FVPHLSI--QANGIGSGNYIGVLSWSPSPHSGQVDLWKIPNYGSSITEATHLAPSV 409  
 420 YPGFNEVIVFMASIPGPNQSGSNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDP 479  
 410 YPGFGEVLVFFMSKMPG--GAYNL-PCLLPQEIYSHLASEQAPTVEAALLHYVDP 465  
 480 TNRNLGEFKLYPGGYLTCVPNSSTGQOOLPLDGVFVFPASWVSFRFYQLKPVGTAGPARGR 539  
 466 TGRNLGEFKAYPDGFLTCVPNGASSGQOOLPINGVFVFSWVSFRFYQLKPVGTASSARGR 525  
 540 LGVRR 544  
 526 LGLRR 530

RESULT 6  
 AAR57091  
 ID AAR57091 standard; protein; 530 AA.  
 XX  
 AC AAR57091;  
 XX  
 XX 27-AUG-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 05-OCT-1994 (first entry)  
 XX  
 DE Small round virus SRSV/KY/89 capsid protein.  
 XX  
 DE pathogen; acute gastroenteritis; food poisoning; seafood contamination;  
 KW diagnostic assay; human calcivirus; small round virus; SRSV; KY89;  
 KW Norwalk virus; capsid protein.  
 XX  
 OS Small round structured virus.  
 XX  
 PN WO9405700-A2.  
 XX  
 PD 17-MAR-1994.  
 XX  
 PF 07-SEP-1993; 93WO-US008447.  
 XX  
 PR 07-SEP-1992; 92US-00941365.  
 XX

```
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI Matson DO, Estes MK, Jiang X, Graham DY;
XX
XX WPI; 1994-101125/12.
XX DR N-PSDB; AAQ56832.
XX
PT DNA from Norwalk and related viruses - used for preparing prods. for use
PT in diagnostic assays, detection and vaccines for Norwalk and related
PT viruses.
XX
XX Example 7; Fig 13a; 156pp; English.
XX
XX The known sequence for Norwalk virus was used to obtain the sequence of
XX other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool
XX from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide
XX cDNA sequence includes part of the polymerase region and the capsid
XX region of the genome; the deduced amino acid sequences are AAR57092 and
XX AAR57091, respectively. Expression of fragments and derivs. of Norwalk-
XX related viruses permits development of diagnostic assays to detect
XX antibodies, antigens, viral genetic material or antivirals. (Updated on
XX 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS
XX field.)
XX
SQ Sequence 530 AA;

Query Match 71.0%; Score 2054; DB 2; Length 530;
Best Local Similarity 71.4%; Pred. No. 4.5e-174;
Matches 389; Conservative 55; Mismatches 85; Indels 16; Gaps 8;

QY 1 MMASKDATPSADGATGAGQLVPEVNTADPIDPVAGSSTALATAGQVNLDPHIINNF 60
DB 1 MMASKDATSSVDGASASQLVPEVNASDPLAMDPAVAGSSTAVATAGQVNPIDPWIINNF 60

QY 61 VOAPQGETTISPNNTPGDLVLDLQGLPHLNPFLSHLSOMYNGWGNMVRVVLGNAFTA 120
DB 61 VOAPQGETTISPNNTPGDLVLDLQGLPHLNPFLSHLSOMYNGWGNMVRVVLGNAFTA 120

QY 121 GKVIICCVPPGQSRSTLSIAQATLFPFHVIADVRTLDPEVEPLEDVRLNLYHNND-TQPTM 179
DB 121 GKIIIVSCIPPGSGSQQLITIAQATLFPFHVIADVRTLDPEVEPLEDVRLNLYHNNDNQTM 180

QY 180 RLLCMLYTLPTLTGGASGGTDSFVAVAGRLTCRGPDPFNFELFVPPVTEOKTRPFTVNPIL 239
DB 181 RLVCMLYTLPTLTGGGTG--DSPVAVAGRLTCRGPDPFNFELFVPPVTEOKTRPFTVNPIL 238

QY 240 KYLSNRIPNPIEGMSLSPDQNTQYQFONGRCTIDGQPLGTTPVSVQLCKEGRITSGQ 299
DB 239 SLSLSNRAPLPIISGMISIPDNVQSVQFONGRCTLDGRLVGTTPVSVLHVAKIRG--TSNG 296

QY 300 RVLNLTLDGSPFMAFAAPAGFPDLGSCDWHIEMSKIPIKNSSTQNNPIVNSVKPNSQQ 359
DB 297 TVINLTLDGTPFHPFEGPAPIGFDFDLGGCDWHIEMSKTQFQHSSTQYDVTI--PDTIS- 352

QY 360 FVPHLSSTLTDENSSGGDIYIGTQNTQYQFONGRCTIDGQPLGTTPVSVQLCKEGRITSGQ 419
DB 353 -VPHLGS- -QANGTIGSGNYIGVLSWSPSPSPSGVQLWKIPNYGSSITEATHLAPSV 409

QY 420 YPPGNEVIVFMASIPGNSQGSNLYVPCLLPQBYIHTFSEQAPIQGEAALLHYVDPD 479
DB 410 YSPGGEVIVFMASIPGNSQGSNLYVPCLLPQBYIHTFSEQAPIQGEAALLHYVDPD 465

QY 480 TNRNLGEFKLYPGGVLTCVPNSSSTGPPQPLDGVFVFASVSRFYQLKPVGTAGPARGR 539
DB 466 TNRNLGEFKLYPGGVLTCVPNSSSTGPPQPLDGVFVFASVSRFYQLKPVGTAGPARGR 525

QY 540 LGVRR 544
DB 526 LGLRR 530

RESULT 7
AAB49700
```

```
ID AAB49700 standard; protein; 545 AA.
XX
AC AAB49700;
XX
XX DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 1.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX WO200079280-A1.
XX
XX 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NTNA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
XX DR N-PSDB; AAF29141.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX
XX Claim 1; Page 40-42; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 545 AA;

Query Match 68.5%; Score 1979; DB 4; Length 545;
Best Local Similarity 67.0%; Pred. No. 2.3e-167;
Matches 368; Conservative 70; Mismatches 101; Indels 10; Gaps 6;

QY 1 MMASKDATPSADGATGAGQLVPEVNTADPIDPVAGSSTALATAGQVNLDPHIINNF 60
DB 1 MMASKDATPNTMDGTSGAGQLVPEANTAEPISEBPVAGATAAATAGQVNMIDPWIINNY 60

QY 61 VOAPQGETTISPNNTPGDLVLDLQGLPHLNPFLSHLSOMYNGWGNMVRVVLGNAFTA 120
DB 61 VOAPQGETTISPNNTPGDLVLDLQGLPHLNPFLSHLSOMYNGWGNMVRVVLGNAFTA 120

QY 121 GKVIICCVPPGQSRSTLSIAQATLFPFHVIADVRTLDPEVEPLEDVRLNLYHNNDQPTMR 180
DB 121 GKIIIVSCIPPGSGSQQLITIAQATLFPFHVIADVRTLDPEVEPLEDVRLNLYHNNDQPTMR 180

QY 181 LLLCMLYTLPTLTGGASGGTDSFVAVAGRLTCRGPDPFNFELFVPPVTEOKTRPFTVNPIL 240
DB 181 LVCMLYTLPLASGSSGTDPPVIAGRLVTCRGPDPFNFELFVPPVTEOKTRPFTVNPIL 240

QY 241 YLSNRIPNPIEGMSLSPDQNTQYQFONGRCTIDGQPLGTTPVSVQLCKEGRITSGQ 298
DB 241 TLSNRVPSLTKSMWVSRDHQGMVQFONGRVTLDGQLGQTTPTASQJCKIRGVSFHANG 300

QY 299 QRVNLTLTDGSPFMAFAAPAGFPDLGSCDWHIEMSKIPIKNSSTQNNPIVNSVKPNSQ 358
DB 301 GNGYNLTLDGSPFMAFAAPAGFPDLGSCDWHIEMSKIPIKNSSTQNNPIVNSVKPNSQ 358

QY 359 QFVPHLSSITLD--ENVSSGGDIYIGTQNTQYQFONGRVTLDGQLGQTTPTASQJCKIRGVSFHANG 416
```

359 AFAPHLGRIQADGSDVSVNTNMIAGWVSVSDGHRGDVDPVPIRYGSTLTERAQLA 418  
 417 PAVYPPGFNEVYFMASIPGNOSGPNL-VPCLLPQYETHFISQAPIQGBAALLHY 475  
 419 PPIYPPGGEALVFMSDFPIAH--GTNGLSVPCCTIPQEFVTHFVNEQAPTRGBAALLHY 476  
 476 VDPDTRNLNIGPKLYPGGYLTCVPSNSTGTGQQLPLDGVFVFSWVSREFYOLKPVGTAGP 535  
 477 LDPDTHRLNIGPKLYPEGFMTCPVNSSTGTGPTLPINGVVFVSWVSREFYOLKPVGTAGP 536  
 536 ARGRLGVR 544  
 537 A-CRUGIR 544

RESULT 8  
 AAB49706  
 ID AAB49706 standard; protein; 540 AA.  
 AC AAB49706;  
 DT 04-APR-2001 (first entry)  
 DE Small round structured virus protein SEQ ID 7.  
 DE Small round structured virus; SRSV; food poisoning.  
 KW Small round structured virus.  
 OS Small round structured virus.  
 PN WO200079280-A1.  
 PD 28-DEC-2000.  
 PF 22-JUN-2000; 2000WO-JP004095.  
 PR 22-JUN-1999; 99JP-00175928.  
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 PA (DENK-) DENKA SEIKEN KK.  
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 XX WPI; 2001-080848/09.  
 XX N-PSDB; AAF29147.  
 XX Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.  
 PS Claim 1; Page 54-57; 84pp; Japanese.  
 CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX Sequence 540 AA;  
 SQ

Query Match 40.5%; Score 1170.5; DB 4; Length 540;  
 Best Local Similarity 45.8%; Pred. No. 3.2e-95;  
 Matches 258; Conservative 85; Mismatches 170; Indels 53; Gaps 16;

QY 1 MMASKDAPSDAGQGVPEVNTADPIPIDPVAGSSTALATAGVNLIDPWNNF 60  
 DB 1 MKWASNDATPSNDGAAG--LVPESSNN-EMAALEPVVGASLAAPVTGQTNIIIDPWIRTF 56  
 QY 61 VQAPQGEFTISPNNTGCVLFDLQGLPHLNPFSLHSOMYGVGNMVRVVLGNAGNFTA 120  
 DB 57 VQAPNGEFTVSPRNSPGEILLVNLDELPELNPYLAHLARMYNGVAGGMEVQVMLAGNFTA 116

QY 121 GKVIICVPPGQSRSLIAQATLFPVHIADVRLDPVEVPLEDVRNVLVH--NNDTOPTM 179  
 DB 117 GKIIFAAAPPYPPVENLSPSQITMPFPHVILIDVRLLEPVLPMPPDVRSTLPHFNQKDEPKM 176  
 QY 180 RLCLMYLPLRTGGASGCTDSFVWAGRVLTCPGDFNPLFLVPPPTVQKTRPFTVPIPL 239  
 DB 177 RLIVAMLYPLRNGS--GDDVTVSCRILTRPSPEFDFYLVPPPTVESKTKPFTLPVLT 234  
 QY 240 KYLSNRIPNPTEGMSLSPDQTONVQFONGRCTIDGQPLGTTTPVSVSOLCKFRGIT--- 296  
 DB 235 GELNSRFPPLSIDEMVTSFNESIVVQPNQGRVTLDELGLLTQLQACNICIRKVKTCOV 294  
 QY 297 -SGQVNL--TELDSPE-MAFAAPAGFPLDSCDWHIEMSKIPNSQNNPIVINS 352  
 DB 295 PSEQHMNLNLTNLNGTQDFTDDVPALGVDFPAGEVGVLSQRNRSNPANRAHDAV 354  
 QY 353 VKPNSQQFVPHLSITLDENVSSGGDIYIGTIOWTSPSPSGGANTNF----- 399  
 DB 355 VATYSDKYTPKGLV-----QIGT--WNT--NDVENOPTKFTPIGLNEVANGHR 399  
 QY 400 ---WKIPDYGSSLAASQLAPAVP--PGFNEVIVVFMASIPGNOSGSPNLVPCLLPOE 454  
 DB 400 FEQWTLPRYSGALTNNNLAPAVAPLPG--ERLLFFRSYVPLKGGFGNP-AIDCSVPOE 456  
 QY 455 YITHFISEQAPIQGEAALLHYVDPDTRNLNIGPKLYPGGYLTCVPSNSTGTGQQLPLDGV 514  
 DB 457 WVQHIFYQESAPSLGDVALVRYVNPDTGRVLFKHLKGGFLTV--SSTSTGPPVVVPANGY 514  
 QY 515 FVFASWVSRYQLKPVGTAGPARGL 540  
 DB 515 FKFDWYNQFYSLAPMGT--GNRRRV 539

RESULT 9  
 AAB49708  
 ID AAB49708 standard; protein; 542 AA.  
 AC AAB49708;  
 DT 04-APR-2001 (first entry)  
 DE Small round structured virus protein SEQ ID 9.  
 DE Small round structured virus; SRSV; food poisoning.  
 KW Small round structured virus.  
 OS Small round structured virus.  
 PN WO200079280-A1.  
 PD 28-DEC-2000.  
 PF 22-JUN-2000; 2000WO-JP004095.  
 PR 22-JUN-1999; 99JP-00175928.  
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 PA (DENK-) DENKA SEIKEN KK.  
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 XX WPI; 2001-080848/09.  
 XX N-PSDB; AAF29149.  
 XX Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.  
 PS Claim 1; Page 59-61; 84pp; Japanese.  
 CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC

CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX  
SQ Sequence 542 AA;

Query Match 40.3%; Score 1165.5; DB 4; Length 542;  
Best Local Similarity 45.3%; Pred. No. 9.1e-95;  
Matches 256; Conservative 89; Mismatches 171; Indels 49; Gaps 16;  
  
QY 1 MMASKDTPSADGATGAGQLVPE-VNTADPTIDPVAGSSALATAGQVNLIDPWINN 59  
DB 1 MMASNDAAPSND---GAASLVPEGIN--ETWLEPVAGASIAAPVAGQTNIIIDPWRTN 55  
  
QY 60 FVQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGAGNAPT 119  
DB 56 FVQAPNGEFTVSPRSPGELLNLNLELGPDLNPLAHLSRMYNGYAGGVEVQVLLAGNAPT 115  
  
QY 120 AGKVIICVPPQFQSGRTLSIAQATLPPHVIADVRLDPEVPLEDVRNVLVYH--NNDTQPT 178  
DB 116 AGKILFAAIPNPLNFDLMDISPAQITMLPHLIVDVRTLEPITPLPDVRNVVYHFNQPPOR 175  
  
QY 179 MELLQWLYPLRTGCGASGCTDSFVAGRVLTCPGDPENFLVLPVTEQKTEPFTVNP 238  
DB 176 MELVAMLYPLRSNGS--GDDVFTVSCRVLTRPTDFEFYLLVPPSVSKTKPFTLPILT 233  
  
QY 239 LKYLNSRIPNPIEGMSLSPDOTQNVQFQNGRGTIDGOLGTPVSVSOLCKFRGRITS- 297  
DB 234 ISELNSRFPPIEQLYAPNETNVVQNGRGTIDGELGQTTQLLSAVCFQGRVTAD 293  
  
QY 298 -----GQRVNLNTELDGSPF--MAFAAPAGFPDILGSCDWHIEMSKIPNSSTQNNPIVTN 351  
DB 294 NGDNWDNLLQTYENGASYDPTDEVPAPLGTQDFSGMLYGV-----LTQDNVNVST 345  
  
QY 352 S-----VKPNSQOFVPHLSSTLDENSSGGDYIGTIQWT--SPPSDGGANTNF- 399  
DB 346 GEAKNAKGIYSTTSKGTFFKIGSIGL-----HSITEHVHPNQOSRFTPVGVAVENTPFPQ 401  
  
QY 400 -WKIPDYGSSLAESAOLAPAVYPPGFNEVIVYFMAISIP--GPNQSGSNLVPCLLPQEVIT 457  
DB 402 QWLVPHYAGSLALNTNLAPAVAPTFFGQLLFRFRSRVPCVQGLQODAFIDCLLQEWVN 461  
  
QY 458 HFISQAPIQGEAALLHYVDPTNRLNGELFKLYPGGYLTCTVNSSTG--PQOLPLDGVF 515  
DB 462 HFYQEAAPSQADVALIRYVNPDTGRTLFEAKLHRSGFITV-----SHTGAYPLVWPNGHF 517  
  
QY 516 VPASVSRFYQLKPVGTAGPARGL 540  
DB 518 RFDVWNVQFYSIAPMGT--GNRRRI 541

RESULT 10  
AAB49704  
ID AAB49704 standard; protein; 539 AA.  
XX  
AC AAB49704;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 5.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
XX  
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX  
DR WPI: 2001-080848/09.  
DR N-PSDB; AAF29145.  
XX  
PT Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX  
XX Claim 1; Page 50-52; 84pp; Japanese.  
XX  
PS This invention relates to a kit for the detection and typing of small  
CC round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX  
SQ Sequence 539 AA;  
  
Query Match 39.8%; Score 1151.5; DB 4; Length 539;  
Best Local Similarity 44.6%; Pred. No. 1.6e-93;  
Matches 254; Conservative 80; Mismatches 175; Indels 61; Gaps 16;  
  
QY 1 MMASKDTPSADGATGAGQLVPE-VNTADPTIDPVAGSSALATAGQVNLIDPWINN 60  
DB 1 MMASNDANPS--DGST--ANLVPEVNN--EVMALPEVGAATAAPVAGQNVDPWIRNF 56  
  
QY 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGAGNAPT 120  
DB 57 VQAPGEFTVSPRNPAGELMSAPLGPDLNPLAHLSRMYNGYAGGVEVQVLLAGNAPT 116  
  
QY 121 GKVIICVPPQFQSGRTLSIAQATLPPHVIADVRLDPEVPLEDVRNVLVYH--NNDTQPTM 179  
DB 117 GKIIIPAAVPPNFPTEGLSPSQVTWPHIIVDVRQLPEVLILPLDVRNMFYHYNQSNDS 176  
  
QY 180 RLICMLYPLRTGCGASGCTDSFVAGRVLTCPGDPENFLVLPVTEQKTEPFTVNP 239  
DB 177 KLIALMYTPLRANNA--GDDVFTVSCRVLTRPSDFDFILVPPVTEGRTEPFTVILT 234  
  
QY 240 KYLNSRIPNPIEGMSLSPDOTQNVQFQNGRGTIDGOLGTPVSVSOLCKFRGRIT--- 296  
DB 235 EEMSNSRFPPIEKLITGTPSSAFVVPQNGRGTIDGVLGTTQLSAVNICTRGVDVTHA 294  
  
QY 297 -SGQRVNLNTELDGSPF--MAFAAPAGFPD--LGSCDWHIEMSKIPNSSTQNNPIVTNS 353  
DB 295 GSHDYTNLNASQWNSNYDPTBEIPAPLGTDPFVGKIQGLMGTQTTREDGSTRAHKA---TV 351  
  
QY 354 KPNQOQFVPHLSSTLDENSSGGDYIGTIQWTSPSPD--SGGANTNF----- 399  
DB 352 STGSHVHTPKL-----GSVQYTTDNTNDFTQGTQTKFTPVGVIQDGNH 395  
  
QY 400 -----WKIPDYGSSLAESAOLAPAVYPPGFNEVIVYFMAISIPGPNQSGSNL--VPCLLP 452  
DB 396 QNEPQWVLPNYSGRTHNVHLAPAVAPTFFGQLLFRSTWPG--CSGYPMNNDCLLP 453  
  
QY 453 QEYITHFISEQAPIQGEAALLHYVDPTNRLNGELFKLYPGGYLTCTVNSSTGPOOL--P 510  
DB 454 QEWQHFCQEAAPQSDVALLRFVNPDTGRTLFECKLHKSQVTV---AHTGPHDLVIP 509  
  
QY 511 LDGVFVFASVSRFYQLKPVGTAGPARGL 540  
DB 510 PNGYFRFDSWVNVQFYTLAPMGNGAGRRAL 539  
  
RESULT 11  
AAB49707  
ID AAB49707 standard; protein; 535 AA.





QY 121 GVIIICVPPGQSRITSLAQATLPHVIADVRLDPEVEPLEDVRNVLVYH--NNDTQPTM 179  
 Db 117 GKIIIFAAIPPNPFDNLSSAAQITMCPHVIDVRQLPEVNLMPDVRRNFFHYNQSGL 176  
 QY 180 RLLCMLYPLRTGGASGGTDSFVAVAGVLTCTGPDNFELFLVPTVEOKTRFTVPNIPL 239  
 Db 177 RLIIAMLYPLRLAN--NSGDDVFTVSCRVLTRSPDPSFNFLVPTVESKTKFTLPIITI 234  
 QY 240 KYLSNSRIPNPIEGMSLSPDQTONVQFONGRCTIDGQPLGTTTPVSVSQLCKFRGRI---- 295  
 Db 235 SEMSNRPFPVPIESLHTSPTENIVVQCQGRVTLTGELMGTTQLLPSRICAIFRGLVTRST 294  
 QY 296 -----TSQORVLN-----LTLDGSPF--MAFAAPAPAGFPDLGSCDWHIEMSKIPN 340  
 Db 295 SRASQADATATPLRLFNYYHVLQDLNLTGTPDPAEDI FGLGTPDPRGKVFVASQRLD 354  
 QY 341 SSTONNPIVTVNSVKNSQOFVPHLSITLTDENVSSGGDYIGTIQWTS--PPSDSG--GAN 396  
 Db 355 STTRAHEA---KVDTTAGRTPKLGSLEISTDSDDFOQNTKFTPVGIGVDNEAE 407  
 QY 397 TNFWKIPDYGSSLAELASQAPAVYPGFNEVIVFYFNASIPGNQSG--SPNLVPCLLPQE 454  
 Db 408 FQWLSLPDYSGQFTHNMNLAPAVAFNFGCEQLLFFRSQLP---SSGGRSNGVLDCLVPQE 464  
 QY 455 YITHFISQAPIQGEAALLHYVDPTNRLNGELFKLPGGYLTCVPSNSSLTGQQQLPLQGV 514  
 Db 465 WQHFYQESAPAQTVQALVRYVNPDTGKVLFEAKLHKLGFMTIANNQGS--PITVPPNGY 522  
 QY 515 VFASVSVSRFYQLKPVGTAGPARGRL 540  
 Db 523 RFESWVNPFYTLAPMGT--GNRRRI 547

## RESULT 13

AAU91272  
 ID AAU91272 standard; protein; 548 AA.

XX AC AAU91272;

DT 18-JUN-2002 (first entry)

DE Norwalk virus associated polynucleotide #1.

KW Norwalk virus; monoclonal antibody; geno group I; geno group II;  
 KW immunological detection; food; viral infection.

OS Norwalk virus.

XX PN JP2002020399-A.

XX PD 23-JAN-2002.

XX PF 10-JUL-2000; 2000JP-00208151.

XX PR 10-JUL-2000; 2000JP-00208151.

XX PA (OSAP ) OSAKA PREFECTURE.

XX PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.

XX PA (IATR ) IATRION LAB INC.

XX DR WPI; 2002-287412/33.

XX PT A monoclonal antibody useful in the immunological detection and diagnosis  
 XX of Norwalk virus infection.

XX PS Disclosure; Page 12-13; 24pp; Japanese.

XX CC The invention describes a monoclonal antibody recognising Norwalk virus,  
 CC a capsid protein of Norwalk virus, or a common antigen epitope on the  
 CC capsid protein molecule of geno group I and geno group II. The antibody  
 CC is useful for immunological detection and quantitative analysis of  
 CC Norwalk virus in foods and the serum of infected patients. This sequence  
 CC represents a Norwalk virus associated protein described in the invention

XX SQ

Sequence 548 AA;

Query Match 39.2%; Score 1134; DB 5; Length 548;

Best Local Similarity 44.5%; Pred No. 6e-92;  
 Matches 248; Conservative 87; Mismatches 180; Indels 42; Gaps 14;

QY 1 MMASKATPDSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLIDPWIINP 60  
 Db 1 MMASRAAPSNDGAAG--LVPEINN--EAMALDPVAGAAATPLTGOQNIIDPWIINP 56

QY 61 VOAPGEBETISPNNTPGDVLFDLOLPHLNPLSHLSOMYNGWGVNMRVVRVLGNAFTA 120  
 Db 57 VOAPGEBETVSPRNSPGBVLNLELGPBEINPYLAHARMYNGYAGGFVQVVLAGNAFTA 116

QY 121 GKVIICVPPGQSRITSLAQATLPHVIADVRLDPEVEPLEDVRNVLVYH--NNDTQPTM 179  
 Db 117 GKIIIFAAIPPNPFDNLSSAAQITMCPHVIDVRQLPEVNLMPDVRRNFFHYNQSGL 176

QY 180 RLLCMLYPLRTGGASGGTDSFVAVAGVLTCTGPDNFELFLVPTVEOKTRFTVPNIPL 239  
 Db 177 RLIIAMLYPLRLAN--NSGDDVFTVSCRVLTRSPDPSFNFLVPTVESKTKFTLPIITI 234

QY 240 KYLSNSRIPNPIEGMSLSPDQTONVQFONGRCTIDGQPLGTTTPVSVSQLCKFRGRI---- 296  
 Db 235 SEMSNRPFPVPIESLHTSPTENIVVQCQGRVTLTGELMGTTQLLPSRICAIFRGLVTRST 294

QY 297 -----TSQORVLN-----LTLDGSPF--MAFAAPAPAGFPDLGSCDWHIEMSKIPN 340  
 Db 295 SRASQADATATPLRLFNYYHVLQDLNLTGTPDPAEDI FGLGTPDPRGKVFVASQRLD 354

QY 341 SSTONNPIVTVNSVKNSQOFVPHLSITLTDENVSSGGDYIGTIQWTS--PPSDSG--GAN 397  
 Db 355 STTRAHEA---KVDTTAGRTPKLGSLEISTDSDDFOQNTKFTPVGIGVDNEAD 408

QY 398 NFWKIPDYGSSLAELASQAPAVYPGFNEVIVFYFNASIPGNQSG--SPNLVPCLLPQE 455  
 Db 409 QWLSLPDYSGQFTHNMNLAPAVAFNFGCEQLLFFRSQLP---SSGGRSNGVLDCLVPQE 465

QY 456 ITHFISQAPIQGEAALLHYVDPTNRLNGELFKLPGGYLTCVPSNSSLTGQQQLPLQGV 515  
 Db 466 WQHFYQESAPAQTVQALVRYVNPDTGKVLFEAKLHKLGFMTIANNQGS--PITVPPNGY 523

QY 516 VFASVSVSRFYQLKPVGT 532

Db 524 RFESWVNPFYTLAPMGT 540

## RESULT 14

AAU91272

ID AAB49709 standard; protein; 550 AA.

XX AC AAB49709;

XX DT 04-APR-2001 (first entry)

XX DE Small round structured virus protein SEQ ID 10.

XX KW Small round structured virus; SRSV; food poisoning.

XX OS Small round structured virus.

XX PN WO200079280-A1.

XX PD 28-DEC-2000.

XX PF 22-JUN-2000; 2000WO-JP004095.

XX PR 22-JUN-1999; 99JP-00175928.

XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 XX (DENK-) DENKA SEIKEN KK.

Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
WPI; 2001-080848/09.  
N-PSDB; AAF29150.

Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

Claim 1; Page 62-64; 84pp; Japanese.

This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

Sequence 550 AA;  
Query Match 38.9%; Score 1124.5; DB 4; Length 550;  
Best Local Similarity 44.2%; Pred. No. 4.3e-91;  
Matches 250; Conservative 80; Mismatches 192; Indels 43; Gaps 14;

1 MMASKDTPSADGATGAGQLVPEVNTADPIDIPVAGSSTALATAGVNLIDPWIINN 60  
1 MKMASNDAAAPSND---GAANLVPEAND-EVMALPEVVGASIAAPVVGQNIIDPWIRENF 56  
61 VQAPOGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSOMYNGVMGNRVVVLGNAFTA 120  
57 VQAPOGEFTVSPNSPGEMLLNLELGPENLPYLHLRMYNGYAGQMVGQVVLGNAFTA 116  
121 GKVIICCVPPGQSRRTLSAQATLPHVIVADVRLDPVEVLEEDVRNVLYH--NNDTQPTM 179  
117 GKIIFAAAPPFPVENISAQITMCCHVIVDRQLEPFLPDINRFFHYNQENTPRM 176  
180 RLICMLYTLRTGASGGTDSFVAGRVLTCPGDFNLFVLPVPTVEQKTRPFTVPNIPL 239  
177 RLVMALYTLPLR---ANGSDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPFTLPL 233  
240 KYLSNRIPIPIEGMSLSPDOTQNVQFNGRGTIDQPLGTTTPVSVSLCKFRGRI--- 295  
234 GELSNRFPAAIDMLYTDPNESVWQFNGRGTIDGTLQGTTLQVPTQICAFGLISQT 293  
296 -----TSQRVLN-----LTELDSGSPF--MAFAAPAPAGFPDLGSCDWHIEMSKIPNS 341  
294 ARAADSTDSQARARNHPLHVQKNLGDQTDPTDDIPAVLGAIDFKGTGFGVASQRDVS 353  
342 STQNNPIVTVNS-----VKPNSQFVPHLSSITLDENVSSGGDYGIT---IQWTSPPSDSG 394  
354 QOEQGHVATRAHEAHIDTTPKVPKLGTL-----IKSGDDDFNTNQPIRFT--PVGMGD 407  
395 ANTWFVKIPDYGSSLAESALAPAVYPGNEVIVFMASIPGPNOSGSPNLVPCLLPQE 454  
408 NNMFQWELPDYSGRLTLNNLAPAVSPFCERILFFRSIVPSAGGYGS--GYIDCLIPQE 466  
455 YITHFISEQAPIGEAALLHYVDPDTRNNGEFLKLYPGGKILTCVPNSSSTGPQOPLDGV 514  
467 WQHFYQEAAPSQAVALLVRYVNDPTGRNIPKALHREGFLTVA--NCGNPIVVPNGY 524  
515 FVPASWSRFPQLKPVGTAGPARG 539  
525 FRFEAWGNQFYLAPMG--SGGRRR 548

RESULT 15  
AAB49710  
ID AAB49710 standard; protein; 541 AA.  
XX  
XX AAB49710;  
XX  
04-APR-2001 (first entry)  
DT

Db 395 GTILIRSDNGHCHDMVGTSPPTTPQWRCRSGSNCCSGHRYPV--PVMNRVTWIV 452  
Qy 432 MASIFGPNGSGSNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPDTNRNLGEFKLYP 491  
Db 453 LSHKSGFSTSTRK-----LPQ-----LNLRW-----LIRFINPDTGRVLPPEARLHK 494  
Qy 492 GGYLTGVPNSSSTGQQQLPDGVFVPASWVSRYQLKPVGTAGPARGRL 540  
Db 495 QGFITVA--HTGDNFIVMPNGYFRFEAWVNOFYSLAPVGT-GKGRRV 540

Search completed: June 1, 2004, 13:45:50  
Job time : 48.7329 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.9611 Seconds  
(without alignments)  
2166.837 Million cell updates/sec

Title: US-09-926-799-4  
Perfect score: 2891  
Sequence: 1 MMASKDTPSADGATGAGQ.....YOLKPVGTAGPARGRLGVR 544

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA\*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2100	72.6	530	US-08-486-049-3	Sequence 3, Appli
2	268	9.3	668	US-09-617-594A-4	Sequence 4, Appli
3	266.5	9.2	626	US-09-590-020-7	Sequence 7, Appli
4	263.5	9.1	669	US-09-617-594A-2	Sequence 2, Appli
5	261.5	9.0	623	US-09-590-020-2	Sequence 2, Appli
6	261.5	9.0	623	US-09-590-020-4	Sequence 4, Appli
7	250.5	8.7	622	US-09-590-020-6	Sequence 6, Appli
8	143.5	5.0	2972	US-09-579-181-2	Sequence 2, Appli
9	143.5	5.0	3118	US-09-579-181-1	Sequence 1, Appli
10	138	4.8	2206	US-07-852-260-2	Sequence 2, Appli
11	138	4.8	2206	US-08-461-503-2	Sequence 2, Appli
12	138	4.8	2206	US-08-465-250-2	Sequence 2, Appli
13	121	4.2	2183	US-08-746-111-5	Sequence 5, Appli
14	119.5	4.1	1234	US-08-317-310A-15	Sequence 15, Appli
15	119.5	4.1	1234	PCT-US95-13041-15	Sequence 15, Appli
16	117.5	4.1	1155	US-08-094-948A-29	Sequence 29, Appli
17	117.5	4.1	1155	PCT-US96-09319-29	Sequence 29, Appli
18	113	3.9	1068	US-08-396-479B-12	Sequence 12, Appli
19	113	3.9	1068	US-08-818-823-12	Sequence 12, Appli
20	111.5	3.9	433	US-09-046-158A-2	Sequence 2, Appli
21	111.5	3.9	897	US-07-960-389-2	Sequence 2, Appli
22	111.5	3.9	2032	US-09-071-035-458	Sequence 458, App
23	111.5	3.9	2032	US-09-071-035-462	Sequence 462, App
24	111.5	3.9	2032	US-09-071-035-466	Sequence 466, App
25	111.5	3.9	2054	US-09-134-000C-6612	Sequence 6612, App
26	111	3.8	1075	PCT-US94-07297-41	Sequence 41, Appli
27	110	3.8	528	US-09-086-663A-82	Sequence 82, Appli

Sequence 2, Appli  
Sequence 80, Appli  
Sequence 26695, A  
Sequence 5503, App  
Sequence 64, Appli  
Sequence 48, Appli  
Sequence 32, Appli  
Sequence 32, Appli  
Sequence 16, Appli  
Sequence 42, Appli  
Sequence 26, Appli  
Sequence 46, Appli  
Sequence 10, Appli  
Sequence 30, Appli  
Sequence 14, Appli  
Sequence 56, Appli  
Sequence 52, Appli

US-08-486-049-3  
Sequence 3, Application US/08486049  
Patent No. 6572862  
GENERAL INFORMATION:  
APPLICANT: BGSB/Mary K  
APPLICANT: Jiang, Xi  
APPLICANT: Graham, David Y  
TITLE OF INVENTION: Methods and Reagents to Detect and  
TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski L.L.P.  
STREET: 801 Pennsylvania Ave., N.W.  
CITY: Washington, D.C.  
STATE:  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,049  
FILING DATE: June 7, 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Davis, Peter  
REGISTRATION NUMBER: 36,119  
REFERENCE/DOCKET NUMBER: 311.023  
TELEPHONE: 202-662-0200  
TELEFAX: 202-662-4643  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-486-049-3

ALIGNMENTS

RESULT 1  
US-08-486-049-3  
Sequence 3, Application US/08486049  
Patent No. 6572862  
GENERAL INFORMATION:  
APPLICANT: BGSB/Mary K  
APPLICANT: Jiang, Xi  
APPLICANT: Graham, David Y  
TITLE OF INVENTION: Methods and Reagents to Detect and  
TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fulbright & Jaworski L.L.P.  
STREET: 801 Pennsylvania Ave., N.W.  
CITY: Washington, D.C.  
STATE:  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,049  
FILING DATE: June 7, 1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Davis, Peter  
REGISTRATION NUMBER: 36,119  
REFERENCE/DOCKET NUMBER: 311.023  
TELEPHONE: 202-662-0200  
TELEFAX: 202-662-4643  
TELEX:  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-486-049-3

Query Match  
Best Local Similarity 72.6%  
Score 2100; DB 4; Length 530;  
Matches 396; Conservative 54; Mismatches 79; Indels 16; Gaps 8;  
Pred. No. 1.5e-187;

QY 1 MMASKDTPSADGATGAGQLVPEVNTADPTPIDVAGSSTALATAGVNLDPWIINNF 60  
DB 1 MMASKDTPSADGATGAGQLVPEVNTADPTPIDVAGSSTALATAGVNLDPWIINNF 60

190 RTGGASGTDSEVAVAGRLVTCGPDENFLFVPTVEQKTRPFTVPNIPKLYLSNSRIPN 249  
296 INPYANDNSGCIIV-TVETKPGDPDFKHLKPG- - - - -SMLTHGSIPS 339  
250 PIEGMSLS - - - - -PDQTONV - - - - -OFQ-NGRCTIDGQPLG-TTPVSUSQCKFRGRI 295  
340 DLIPKSSSLWGNRHSIDITDFVIKPFVQANRHFDFNQETAGWSTP - - - - -RFR-PI 391  
296 TSGQRLVNLTELDGSPFFMAFAP - - - - -APAGPDLGSCDWHIEMSKIPNSST-QNNPI 348  
392 T - - - - -ITVSEKGGKLGIGVATDSIVPGIPDGMWD - - - - -TTIPEKLIAPAGDYA 436  
349 VTNVSKPNSQFVPHLSITLDENVSSGDYI-GTIQ--WTSPPSDSGGANTNFWKIPDY 405  
437 ITNGGNDITTAADYDGSIIKNNINFGMYICGALQRAW - - - - -GDKKISNTAF - - - - -I 487  
406 GSSLAELASOLAPA 418  
488 TTAIREGNSIKPS 500  
RESULT 3  
US-09-590-020-7  
; Sequence 7, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vlinis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590, 020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 626  
; TYPE: PRF  
; ORGANISM: Feline calicivirus  
US-09-590-020-7  
Query Match 9.2%; Score 266.5; DB 4; Length 626;  
Best Local Similarity 33.8%; Pred. No. 4.7e-16;  
Matches 74; Conservative 32; Mismatches 88; Indels 25; Gaps 7;  
QY 13 DGATGAGQLVPEVNT-ADPPIPDVAGSSTALATAGQVNLIDPWIINNFOAQOGEFTI- 70  
DB 82 DGSITA - - - - -PEQGTWVGVIAEPSAQWSTAAADWATGKSVSEW - - - - -EAFFSFH 128  
QY 71 - - - - -SPNNTPGDVLFDLQGLHNPFLSHLSQMYNGVGNMVRVVLVAGNAFTAGKVII 125  
DB 129 TSNVMSSTSETQKILFKQSLGSLPILNLYLAKLYVAMWSGSIEVRSISGSGVFGKLA 188  
QY 126 CCVPPGFGQS-RTLSIAQATLPHVIADVRLDPEVEPLEDRVNLVYHNNDTQPTMRLC 184  
DB 189 IVVPPGVDPVQSTMLQ - - - - -YHVLFDARQVEPVIFCLPDLRSTLYHLMSDITTSVIM 245  
QY 185 LYTFLRTGGASGTDSEVAVAGRLVTCGPDENFLFVPT 223  
DB 246 VYNDLINPYANDTNSGCCIV-TVETKPGDPDFKHLKPP 283

RESULT 4  
US-09-617-594A-2  
; Sequence 2, Application US/09617594A  
; Patent No. 6541458  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT

61 VOAPQGEFTISNNTPGDVLFDLQGLHNPFLSHLSQMYNGVGNMVRVVLVAGNAFTA 120  
61 VOAPQGEFTISNNTPGDVLFDLQGLHNPFLSHLSQMYNGVGNMVRVVLVAGNAFTA 120  
121 GKVIICCVPPGQSRTLSIAQATLPHVIADVRLDPEVEPLEDRVNLVYHNNDT-QPTM 179  
121 GKIIIVSCIPPPGGSNHLITIAQATLPHVIADVRLDPEVEPLEDRVNLVYHNNDT-QPTM 180  
180 RLICMLYTEPLRTGGASGTDSEVAVAGRLVTCGPDENFLFVPTVEQKTRPFTVPNIP 239  
181 RLVCMLYTEPLRTGGGTG--DSFVAVAGRVMTCPSPDNFLFVPTVEQKTRPFTVPNIP 238  
240 KYLSNSRIINPEIGMSLSDQTONVQFQNGRCTIDGQPLGTTTPVSUSQCKFRGRI 299  
239 SLSNSRAELPISMGISPDNVQSVQFQNGRCTIDGRLVGTTPVLSHVAKIRG--TSNG 296  
300 RVNLNLTDLGSPFMAFAAPAGPDLGSCDWHIEMSKIPNSSTQNNPIVTVNSVKPNSQ 359  
297 TVINLTDLGTPPHPEGPAPIGPDLGGCDWHIINNTOFGHSSQTYDVT--PDT-- 351  
360 FVPHLSITLDENVSSGDYIYGTIOWTSPPSDSGGANTNFWKIPDYGSSLAELASOLAPAV 419  
352 FVPHLSI--QANGIGSGNYGVLSWISPPSHPSGSDWLMKIPNGSSITEATHLAPSV 409  
420 YPGENEVIVFVWASIPGNQSGPNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDP 479  
410 YPGFGEVLVFFNKGMPG--GAYNL-PCLLPQEIYTHLASEQAPTVGEAALLHYVDP 465  
480 TNRNLGSKYLPQGYLTCVNSNSSTGQQLPDCGVFVAFASWRSFYOLKPVGTAGPAR 539  
466 TGRNLGSKYLPQGYLTCVNSNSSTGQQLPDCGVFVAFASWRSFYOLKPVGTAGPAR 525  
540 LGVR 544  
526 LGLR 530  
RESULT 2  
US-09-617-594A-4  
; Sequence 4, Application US/09617594A  
; Patent No. 6541458  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.1  
; CURRENT APPLICATION NUMBER: US/09/617,594A  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; TYPE: PRF  
; ORGANISM: Feline calicivirus  
US-09-617-594A-4  
Query Match 9.3%; Score 268; DB 4; Length 668;  
Best Local Similarity 27.6%; Pred. No. 3.8e-16;  
Matches 103; Conservative 54; Mismatches 140; Indels 76; Gaps 18;  
QY 71 SPNNTPGDVLFDLQGLHNPFLSHLSQMYNGVGNMVRVVLVAGNAFTAGKVIICVPP 130  
DB 179 STSETQKILFKQSLGSLPILNLYLAKLYVAMWSGSIEVRSISGSGVFGKLAIVVPP 238  
QY 131 GFGS-RTLSIAQATLPHVIADVRLDPEVEPLEDRVNLVYHNNDTQPTMRLCMLYTP 189  
DB 239 GIDPVQSTMLQ - - - - -YHVLFDARQVEPVIFITPDLRNSLYHLMSDITTSVIMYNDL 295

```
; FILE REFERENCE: 454313-3151.1
; CURRENT APPLICATION NUMBER: US/09/617,594A
; CURRENT FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-2

Query Match          9.1%; Score 263.5; DB 4; Length 669;
Best Local Similarity 32.7%; Pred. No. 1.e-15;
Matches 72; Conservative 33; Mismatches 92; Indels 23; Gaps 7;

Qy 12 ADGATGAGOLPEVNT-ADPIPDPVAGSSTALATAGQVNLIDPWIINNFOVQOGCEFTI 70
Db 125 ADGSSII--TPEQGLVGGVIAEBSAQWATAAATGKSDSEW-----ESFESF 173

Qy 71 -----SPNTPGDVFLDLQGLPHLNPFLSHLSQMYNGVGNMRVVRVVLGNAGNAFTAGKVI 124
Db 174 HTSVNMTSETQKILFKQSLGPLNPLYLHLSKLYVWAGSVDVRFSGSGVFGGKLA 233

Qy 125 ICCVPGFQS-RTLSIAQATLFPVHIADVRTLDPEVPLEDRVNLVHNNDTQPTMLLC 183
Db 234 AIVWPGVDVFOQSTSMQLQ---YPHVLFDARQVEPVIFSIDRLSTLYHLMSDITDTTSLVI 290

Qy 184 MLYTLRTGGAGGTDTSFVAGRVLTCPGDPDNFLFLVPP 223
Db 291 MYNDLNPYANDSNSGCIV-TVETKPGDPFKFHLKPP 329

RESULT 5
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-2

Query Match          9.0%; Score 261.5; DB 4; Length 623;
Best Local Similarity 38.3%; Pred. No. 1.4e-15;
Matches 59; Conservative 27; Mismatches 63; Indels 5; Gaps 3;

Qy 71 SPNTPGDVFLDLQGLPHLNPFLSHLSQMYNGVGNMRVVRVVLGNAGNAFTAGKVIICVPP 130
Db 134 STSETQKILFKQSLGPLNPLYLHLSKLYVWAGSVEVRFSGSGVFGGKLAIVVPP 193

Qy 131 GFQS-RTLSIAQATLFPVHIADVRTLDPEVPLEDRVNLVHNNDTQPTMLLCMLYTPL 189
Db 194 GIEPVQSTSMQLQ---YPHVLFDARQVEPVIFAIPOLRSNLYHLMSDITDTTSLVIMVYNDL 250

RESULT 6
US-09-590-020-4
; Sequence 4, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match          9.0%; Score 261.5; DB 4; Length 623;
Best Local Similarity 38.3%; Pred. No. 1.4e-15;
Matches 59; Conservative 27; Mismatches 63; Indels 5; Gaps 3;

Qy 71 SPNTPGDVFLDLQGLPHLNPFLSHLSQMYNGVGNMRVVRVVLGNAGNAFTAGKVIICVPP 130
Db 134 STSETQKILFKQSLGPLNPLYLHLSKLYVWAGSVEVRFSGSGVFGGKLAIVVPP 193

Qy 131 GFQS-RTLSIAQATLFPVHIADVRTLDPEVPLEDRVNLVHNNDTQPTMLLCMLYTPL 189
Db 194 GIEPVQSTSMQLQ---YPHVLFDARQVEPVIFAIPOLRSNLYHLMSDITDTTSLVIMVYNDL 250

RESULT 7
US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-6

Query Match          8.7%; Score 250.5; DB 4; Length 622;
Best Local Similarity 37.7%; Pred. No. 1.5e-14;
Matches 58; Conservative 27; Mismatches 64; Indels 5; Gaps 3;

Qy 71 SPNTPGDVFLDLQGLPHLNPFLSHLSQMYNGVGNMRVVRVVLGNAGNAFTAGKVIICVPP 130
```

Db 134 STSETQKILFKQSLGFLNLYLHLAKLYVALAGSEVFRFSISGSGVFGKLAIVVPP 193  
Qy 131 GFQS-RTLSIAQATLFPFHVIADVRLDOPVEVPLEDRNVLVHNNDQPTWRLLCMLYTPL 189  
Db 194 GIEPVQSTSMQLQ--YPHVLFDAQVEVIEFAIPDLRSNLYHLMSDITDTTSLVIMVYNDL 250  
Qy 190 RTFGASGGTDSFVAVAGRLTCGPDFFNLFVPP 223  
Db 251 INPYANDTSSGCIIV-TVETKPGDPFKHLLKPP 283  
RESULT 8  
US-09-579-181-2  
; Sequence 2, Application US/09579181  
; Patent No. 6365372  
; GENERAL INFORMATION:  
; APPLICANT: Chrivia, John  
; APPLICANT: Yaciuk, Peter  
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)  
; FILE REFERENCE: 16153-4247  
; CURRENT APPLICATION NUMBER: US/09/579,181  
; CURRENT FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/136,620  
; PRIOR FILING DATE: 1999-05-27  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2  
; LENGTH: 2972  
; TYPE: PRT  
; ORGANISM: Human  
US-09-579-181-2  
Query Match 5.0%; Score 143.5; DB 4; Length 2972;  
Best Local Similarity 22.1%; Pred. No. 0.0019;  
Matches 124; Conservative 51; Mismatches 226; Indels 159; Gaps 22;  
Qy 19 GOLVPEVNT-----ADPIDPVGSS--TALATAGQVNLDPWLNINNVQA--POGETTISP 72  
Db 809 GRTVVVNNPRAPLGPVPRPPGPELSAQPTGFPVQVLP--ASLMVSASAPGPPILPA 866  
Qy 73 NNTGCDVLFDLQGLPHLNPFLSHLSQMYNGWGNMR-----VRVVL 113  
Db 867 SRPPGPVLL-----PPQPNSGSLPQVLPSPGLVSGTSRPPPTLSLKETPPAPVRLSP 921  
Qy 114 AGNAFTAGKVIICCVPPGF-----QSRTLSIAQATLFPFHVIADVRLDPEVPLED 164  
Db 922 APPGSSLLKPLTVPPGCTPPAAATTTTATATATTAVP-----APTAPQRLILSP 976  
Qy 165 VRNVLYHNNDQPTWRLLCMLYPLRTGASGGTDSFVAVAGRLTCGPDFFNLF----- 220  
Db 977 DMQARLPSEVVISIGQLASLAQRVANAGGSKPL-TFOIQGNKLTLTGAQVRLAVGQPR 1035  
Qy 221 ---VPPT-----VEQTRPFTVNIPL-----KYLNSRIP 248  
Db 1036 PLQMPPTMNNTVGVVKIVVRQAPRDGLTPVPLAPAPRPPSSGLPAVLNPRPTLTGRLP 1095  
Qy 249 NPIEGMSLSPDQTNQVQNGRCITDGOPLGTTTPVSVSQLCKFRGRTTSGORVLNLTELD 308  
Db 1096 TPTLTGTARAP-----MP---TPTLVRLPK-----LV 1119  
Qy 309 GSPFMAFAAPAPAGPDLGSCDWHIEMS-----KIPNSSTQNNPIVTVNSVKPNSQQ 359  
Db 1120 HSPSPVVSASAPGAAPLTITSSPLHVPSSLPGPASSPMPINPSSPLASVSVTVSVPLSSS 1179  
Qy 360 F-----VPHLSITLDENVSSGGYIGTIQWTSF-----PSDSCGANTNFWK 401  
Db 1180 LPTISVPTTLPAAPASAPLTIPISAPLTVSASGALLT-SVTPLAPVVPAAFGPPSLQ--- 1235  
Qy 402 IPDYGSSLABASQALAPVY-----PGFNEVIVVFMASIPGPNOSGSPNLVPCILLPOE 454  
Db 1236 -PSGASPSASALTGLATAPSLSSSQTPGHPHLLAPTSSHVPGNLSTVAPACSPVLVPAS 1294

Qy 455 YI-THFISEQAPIQGEAALL 473  
Db 1295 ALASPPPSAPNPAPAQAALL 1314  
RESULT 9  
US-09-579-181-1  
; Sequence 1, Application US/09579181  
; Patent No. 6365372  
; GENERAL INFORMATION:  
; APPLICANT: Chrivia, John  
; APPLICANT: Yaciuk, Peter  
; TITLE OF INVENTION: SNF2 Related CBP Activator Protein (SRCAP)  
; FILE REFERENCE: 16153-4247  
; CURRENT APPLICATION NUMBER: US/09/579,181  
; CURRENT FILING DATE: 2000-05-25  
; PRIOR APPLICATION NUMBER: 60/136,620  
; PRIOR FILING DATE: 1999-05-27  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1  
; LENGTH: 3118  
; TYPE: PRT  
; ORGANISM: Human  
US-09-579-181-1  
Query Match 5.0%; Score 143.5; DB 4; Length 3118;  
Best Local Similarity 22.1%; Pred. No. 0.0021;  
Matches 124; Conservative 51; Mismatches 226; Indels 159; Gaps 22;  
Qy 19 GOLVPEVNT-----ADPIDPVGSS--TALATAGQVNLDPWLNINNVQA--POGETTISP 72  
Db 955 GRTVVVNNPRAPLGPVPRPPGPELSAQPTGFPVQVLP--ASLMVSASAPGPPILPA 1012  
Qy 73 NNTGCDVLFDLQGLPHLNPFLSHLSQMYNGWGNMR-----VRVVL 113  
Db 1013 SRPPGPVLL-----PPQPNSGSLPQVLPSPGLVSGTSRPPPTLSLKETPPAPVRLSP 1067  
Qy 114 AGNAFTAGKVIICCVPPGF-----QSRTLSIAQATLFPFHVIADVRLDPEVPLED 164  
Db 1068 APPGSSLLKPLTVPPGCTPPAAATTTTATATATTAVP-----APTAPQRLILSP 1122  
Qy 165 VRNVLYHNNDQPTWRLLCMLYPLRTGASGGTDSFVAVAGRLTCGPDFFNLF----- 220  
Db 1123 DMQARLPSEVVISIGQLASLAQRVANAGGSKPL-TFOIQGNKLTLTGAQVRLAVGQPR 1181  
Qy 221 ---VPPT-----VEQTRPFTVNIPL-----KYLNSRIP 248  
Db 1182 PLQMPPTMNNTVGVVKIVVRQAPRDGLTPVPLAPAPRPPSSGLPAVLNPRPTLTGRLP 1241  
Qy 249 NPIEGMSLSPDQTNQVQNGRCITDGOPLGTTTPVSVSQLCKFRGRTTSGORVLNLTELD 308  
Db 1242 TPTLTGTARAP-----MP---TPTLVRLPK-----LV 1265  
Qy 309 GSPFMAFAAPAPAGPDLGSCDWHIEMS-----KIPNSSTQNNPIVTVNSVKPNSQQ 359  
Db 1266 HSPSPVVSASAPGAAPLTITSSPLHVPSSLPGPASSPMPINPSSPLASVSVTVSVPLSSS 1325  
Qy 360 F-----VPHLSITLDENVSSGGYIGTIQWTSF-----PSDSCGANTNFWK 401  
Db 1326 LPTISVPTTLPAAPASAPLTIPISAPLTVSASGALLT-SVTPLAPVVPAAFGPPSLQ--- 1381  
Qy 402 IPDYGSSLABASQALAPVY-----PGFNEVIVVFMASIPGPNOSGSPNLVPCILLPOE 454  
Db 1382 -PSGASPSASALTGLATAPSLSSSQTPGHPHLLAPTSSHVPGNLSTVAPACSPVLVPAS 1440  
Qy 455 YI-THFISEQAPIQGEAALL 473  
Db 1441 ALASPPPSAPNPAPAQAALL 1460  
RESULT 10  
US-07-852-260-2



; Sequence 2, Application US/07852260  
; Patent No. 5525715  
; GENERAL INFORMATION:  
; APPLICANT: Racaniello, Vincent  
; APPLICANT: Tatem, Joanne M.  
; APPLICANT: Weeks-Levy, Carolyn L.  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM  
; TITLE OF INVENTION: CDNA  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/852,260  
; FILING DATE: 19920619  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 977-9550  
; TELEFAX: (212) 664-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2206 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-852-260-2

Query Match 4.8%; Score 138; DB 1; Length 2206;  
Best Local Similarity 24.2%; Pred. No. 0.0038;  
Matches 61; Conservative 31; Mismatches 110; Indels 50; Gaps 10;  
  
QY 22 VPEVNTADPIPDVAGSSTALATAGQVNLDPWIINNFVQAQG-----EFTISPNTP 76  
Db 365 IPEFVTPPIDIPGVKNMELAE-----IDTMIPLNLESTKRTNMDMYRVTLSDSADL 418  
  
QY 77 GDVLFDLQLGPHLNPFLSH-----LSOMYNGWGNMVRVVLGNAGTAKVVICCVPPG 131  
Db 419 SQPILCLSLSPAFDPRLSHTMLGEVLNYYTHWAGSLKFTFLFCGSMMATGKILVAYAPG 478  
  
QY 132 FQSRTLSIAQATLPHVIADVTLDPVEPLEDVNRVLYHNNDTPTMR---LLCMLY-- 186  
Db 479 AQPPT-SRKEAMLGTHVINDLGLOQSSCTMVVWVWISNVTYRQT-TQDSFTEGGYISMFYQT 536  
  
QY 187 ---TPLRTGGASGDTSFVAGRVLTCPGDPNFELFVPTVEQKTRPTVTNIPKYL 243  
Db 537 RIWVPLST-----PKSMSMLGFVSAC--NDFSRLRLDRTT-----HIS 572  
  
QY 244 NSRIENPIEGMS 255  
Db 573 QSALPQGIEDLT 584

RESULT 11  
US-07-852-260-2  
; Sequence 2, Application US/08461503  
; Patent No. 5834302  
; GENERAL INFORMATION:  
; APPLICANT: Racaniello, Vincent  
; APPLICANT: Tatem, Joanne M.

; APPLICANT: Weeks-Levy, Carolyn L.  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES  
; TITLE OF INVENTION: FROM CDNA  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,503  
; FILING DATE: 5-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2206 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-461-503-2

Query Match 4.8%; Score 138; DB 2; Length 2206;  
Best Local Similarity 24.2%; Pred. No. 0.0038;  
Matches 61; Conservative 31; Mismatches 110; Indels 50; Gaps 10;  
  
QY 22 VPEVNTADPIPDVAGSSTALATAGQVNLDPWIINNFVQAQG-----EFTISPNTP 76  
Db 365 IPEFVTPPIDIPGVKNMELAE-----IDTMIPLNLESTKRTNMDMYRVTLSDSADL 418  
  
QY 77 GDVLFDLQLGPHLNPFLSH-----LSOMYNGWGNMVRVVLGNAGTAKVVICCVPPG 131  
Db 419 SQPILCLSLSPAFDPRLSHTMLGEVLNYYTHWAGSLKFTFLFCGSMMATGKILVAYAPG 478  
  
QY 132 FQSRTLSIAQATLPHVIADVTLDPVEPLEDVNRVLYHNNDTPTMR---LLCMLY-- 186  
Db 479 AQPPT-SRKEAMLGTHVINDLGLOQSSCTMVVWVWISNVTYRQT-TQDSFTEGGYISMFYQT 536  
  
QY 187 ---TPLRTGGASGDTSFVAGRVLTCPGDPNFELFVPTVEQKTRPTVTNIPKYL 243  
Db 537 RIWVPLST-----PKSMSMLGFVSAC--NDFSRLRLDRTT-----HIS 572  
  
QY 244 NSRIENPIEGMS 255  
Db 573 QSALPQGIEDLT 584

RESULT 12  
US-08-465-250-2  
; Sequence 2, Application US/08465250  
; Patent No. 6136570  
; GENERAL INFORMATION:  
; APPLICANT: Racaniello, Vincent  
; APPLICANT: Tatem, Joanne M.  
; APPLICANT: Weeks-Levy, Carolyn L.  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM  
; TITLE OF INVENTION: CDNA  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:

```

;
; COUNTRY: United States of America
; ZIP: 94104
;

```

COMPUTER: IBM PC compatible

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60  
61  
62  
63  
64  
65  
66  
67  
68  
69  
70  
71  
72  
73  
74  
75  
76  
77  
78  
79  
80  
81  
82  
83  
84  
85  
86  
87  
88  
89  
90  
91  
92  
93  
94  
95  
96  
97  
98  
99  
100  
101  
102  
103  
104  
105  
106  
107  
108  
109  
110  
111  
112  
113  
114  
115  
116  
117  
118  
119  
120  
121  
122  
123  
124  
125  
126  
127  
128  
129  
130  
131  
132  
133  
134  
135  
136  
137  
138  
139  
140  
141  
142  
143  
144  
145  
146  
147  
148  
149  
150  
151  
152  
153  
154  
155  
156  
157  
158  
159  
160  
161  
162  
163  
164  
165  
166  
167  
168  
169  
170  
171  
172  
173  
174  
175  
176  
177  
178  
179  
180  
181  
182  
183  
184  
185  
186  
187  
188  
189  
190  
191  
192  
193  
194  
195  
196  
197  
198  
199  
200  
201  
202  
203  
204  
205  
206  
207  
208  
209  
210  
211  
212  
213  
214  
215  
216  
217  
218  
219  
220  
221  
222  
223  
224  
225  
226  
227  
228  
229  
230  
231  
232  
233  
234  
235  
236  
237  
238  
239  
240  
241  
242  
243  
244  
245  
246  
247  
248  
249  
250  
251  
252  
253  
254  
255  
256  
257  
258  
259  
260  
261  
262  
263  
264  
265  
266  
267  
268  
269  
270  
271  
272  
273  
274  
275  
276  
277  
278  
279  
280  
281  
282  
283  
284  
285  
286  
287  
288  
289  
290  
291  
292  
293  
294  
295  
296  
297  
298  
299  
300  
301  
302  
303  
304  
305  
306  
307  
308  
309  
310  
311  
312  
313  
314  
315  
316  
317  
318  
319  
320  
321  
322  
323  
324  
325  
326  
327  
328  
329  
330  
331  
332  
333  
334  
335  
336  
337  
338  
339  
340  
341  
342  
343  
344  
345  
346  
347  
348  
349  
350  
351  
352  
353  
354  
355  
356  
357  
358  
359  
360  
361  
362  
363  
364  
365  
366  
367  
368  
369  
370  
371  
372  
373  
374  
375  
376  
377  
378  
379  
380  
381  
382  
383  
384  
385  
386  
387  
388  
389  
390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405  
406  
407  
408  
409  
410  
411  
412  
413  
414  
415  
416  
417  
418  
419  
420  
421  
422  
423  
424  
425  
426  
427  
428  
429  
430  
431  
432  
433  
434  
435  
436  
437  
438  
439  
440  
441  
442  
443  
444  
445  
446  
447  
448  
449  
450  
451  
452  
453  
454  
455  
456  
457  
458  
459  
460  
461  
462  
463  
464  
465  
466  
467  
468  
469  
470  
471  
472  
473  
474  
475  
476  
477  
478  
479  
480  
481  
482  
483  
484  
485  
486  
487  
488  
489  
490  
491  
492  
493  
494  
495  
496  
497  
498  
499  
500  
501  
502  
503  
504  
505  
506  
507  
508  
509  
510  
511  
512  
513  
514  
515  
516  
517  
518  
519  
520  
521  
522  
523  
524  
525  
526  
527  
528  
529  
530  
531  
532  
533  
534  
535  
536  
537  
538  
539  
540  
541  
542  
543  
544  
545  
546  
547  
548  
549  
550  
551  
552  
553  
554  
555  
556  
557  
558  
559  
560  
561  
562  
563  
564  
565  
566  
567  
568  
569  
570  
571  
572  
573  
574  
575  
576  
577  
578  
579  
580  
581  
582  
583  
584  
585  
586  
587  
588  
589  
590  
591  
592  
593  
594  
595  
596  
597  
598  
599  
600  
601  
602  
603  
604  
605  
606  
607  
608  
609  
610  
611  
612  
613  
614  
615  
616  
617  
618  
619  
620  
621  
622  
623  
624  
625  
626  
627  
628  
629  
630  
631  
632  
633  
634  
635  
636  
637  
638  
639  
640  
641  
642  
643  
644  
645  
646  
647  
648  
649  
650  
651  
652  
653  
654  
655  
656  
657  
658  
659  
660  
661  
662  
663  
664  
665  
666  
667  
668  
669  
670  
671  
672  
673  
674  
675  
676  
677  
678  
679  
680  
681  
682  
683  
684  
685  
686  
687  
688  
689  
690  
691  
692  
693  
694  
695  
696  
697  
698  
699  
700  
701  
702  
703  
704  
705  
706  
707  
708  
709  
710  
711  
712  
713  
714  
715  
716  
717  
718  
719  
720  
721  
722  
723  
724  
725  
726  
727  
728  
729  
730  
731  
732  
733  
734  
735  
736  
737  
738  
739  
740  
741  
742  
743  
744  
745  
746  
747  
748  
749  
750  
751  
752  
753  
754  
755  
756  
757  
758  
759  
760  
761  
762  
763  
764  
765  
766  
767  
768  
769  
770  
771  
772  
773  
774  
775  
776  
777  
778  
779  
780  
781  
782  
783  
784  
785  
786  
787  
788  
789  
790  
791  
792  
793  
794  
795  
796  
797  
798  
799  
800  
801  
802  
803  
804  
805  
806  
807  
808  
809  
810  
811  
812  
813  
814  
815  
816  
817  
818  
819  
820  
821  
822  
823  
824  
825  
826  
827  
828  
829  
830  
831  
832  
833  
834  
835  
836  
837  
838  
839  
840  
84

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION FORM:
; APPLICATION NUMBER: US/08/317,310A
; FILING DATE: 03-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: JDP-022
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1234 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; US-08-317-310A-15

```

```

Query Match          4.1%; Score 119.5; DB 2; Length 1234;
Best Local Similarity 20.5%; Pred. No. 0.079;
Matches 98; Conservative 68; Mismatches 140; Indels 171; Gaps 29;

QY 156 DPEVPLEDRVNLVYHNNDTQTMRLLCMLYPLRTGASGCTDSFVWAGRVLTCPGPDF 215
Db 273 NPISVPLR-----RHLLNPPPSQ---VGLTRRSRTESITATSPASWVGK-----PGS--- 318
QY 216 NLFVLVPTVEOK---TRPFTVNIPL-----KYLNSRIPNPIEGMSLSPDQTO 262
Db 319 ---FRVRASDGEGTMSRPASVDGSPVSPSTNRTHAHRHSGSLRHPPL-----NHSR 368
QY 263 NVQFONGRCTIDGQPLGTTTPVSQSL-----CKFRGRITSGQRLNLTDLGSPF 312
Db 369 SIPMPSRCS---PSATSPVLSLSSSTSGHGSTDCLPFRSSA-----SVSGS--- 414
QY 313 MAFAPAPAGF---PDLGS---CDWHIEMSKIPINSSTQNNPIVTNSVKPNSQOFVPHLSI 367
Db 415 ----PSDGGFISSEYSGSPCDFRSSFRSVTPDSLGHTPPARGEEELSNYICMGKGAS 469
QY 368 TLDEN-----VSSGGD---YI-GTIQWTSF--PSD--SGGAN-----T 397
Db 470 TLAAPNGHYILSRGGNGHRYIPGANLGTSPALPGDEAAGADLNRFRKRTHSAGTSPTI 529
QY 398 NFWKIPDYS--SLAEASOLAPAVYPPGFNEVIVFMAISIPGNOSGSPNLPCLLPQBY 455
Db 530 SHQKTPSQSSVASIEEYTEMPPAAYPPG-----GGSGRLPGYRHSAP 572
QY 456 I-THFISEQAPIQGEAALLHYVD-----PDTNRNL-----GEFKLYP----- 491
Db 573 VTHSYPE-----EGLEMHLLERRGGHRRPDTN-NLHTDDGYMPMSPGVAPVPSNRKGN 625
QY 492 GYLTCVPSNNSST-----GQQLPLDGVFVFWASWVRFYQLKPVGTAGPARG 538
Db 626 GDMPMSPKSVSAPOQIINPIRRHPQVRDPNG-----YMMWSPSGSCSPDIG 672

```

## RESULT 15

```

PCT-US95-13041-15
; Sequence 15, Application PC/TUS9513041
; GENERAL INFORMATION:
; APPLICANT: WHITE, Morris F.
; APPLICANT: SUN, Xiao Jian
; APPLICANT: PIERCE, Jacalyn H.
; TITLE OF INVENTION: THE IRS FAMILY OF GENES
; NUMBER OF SEQUENCES: 63
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 State Street, Suite 510
; CITY: Boston
; STATE: Massachusetts

```

```

; COUNTRY: USA
; ZIP: 02109-1875
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/13041
; FILING DATE: Herewith
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/317,310
; FILING DATE: 03-OCT-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Louis Myers
; REGISTRATION NUMBER: 35,965
; REFERENCE/DOCKET NUMBER: JDP-022PC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617)227-7400
; TELEFAX: (617)227-5941
; INFORMATION FOR SEQ ID NO: 15:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1234 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; PCT-US95-13041-15

```

```

Query Match          4.1%; Score 119.5; DB 5; Length 1234;
Best Local Similarity 20.5%; Pred. No. 0.079;
Matches 98; Conservative 68; Mismatches 140; Indels 171; Gaps 29;

QY 156 DPEVPLEDRVNLVYHNNDTQTMRLLCMLYPLRTGASGCTDSFVWAGRVLTCPGPDF 215
Db 273 NPISVPLR-----RHLLNPPPSQ---VGLTRRSRTESITATSPASWVGK-----PGS--- 318
QY 216 NLFVLVPTVEOK---TRPFTVNIPL-----KYLNSRIPNPIEGMSLSPDQTO 262
Db 319 ---FRVRASDGEGTMSRPASVDGSPVSPSTNRTHAHRHSGSLRHPPL-----NHSR 368
QY 263 NVQFONGRCTIDGQPLGTTTPVSQSL-----CKFRGRITSGQRLNLTDLGSPF 312
Db 369 SIPMPSRCS---PSATSPVLSLSSSTSGHGSTDCLPFRSSA-----SVSGS--- 414
QY 313 MAFAPAPAGF---PDLGS---CDWHIEMSKIPINSSTQNNPIVTNSVKPNSQOFVPHLSI 367
Db 415 ----PSDGGFISSEYSGSPCDFRSSFRSVTPDSLGHTPPARGEEELSNYICMGKGAS 469
QY 368 TLDEN-----VSSGGD---YI-GTIQWTSF--PSD--SGGAN-----T 397
Db 470 TLAAPNGHYILSRGGNGHRYIPGANLGTSPALPGDEAAGADLNRFRKRTHSAGTSPTI 529
QY 398 NFWKIPDYS--SLAEASOLAPAVYPPGFNEVIVFMAISIPGNOSGSPNLPCLLPQBY 455
Db 530 SHQKTPSQSSVASIEEYTEMPPAAYPPG-----GGSGRLPGYRHSAP 572
QY 456 I-THFISEQAPIQGEAALLHYVD-----PDTNRNL-----GEFKLYP----- 491
Db 573 VTHSYPE-----EGLEMHLLERRGGHRRPDTN-NLHTDDGYMPMSPGVAPVPSNRKGN 625
QY 492 GYLTCVPSNNSST-----GQQLPLDGVFVFWASWVRFYQLKPVGTAGPARG 538
Db 626 GDMPMSPKSVSAPOQIINPIRRHPQVRDPNG-----YMMWSPSGSCSPDIG 672

```

Search completed: June 1, 2004, 13:57:57  
Job time : 13.9611 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.7758 Seconds  
(without alignments)  
4368.312 Million cell updates/sec

Title: US-09-926-799-4  
Perfect score: 2891  
Sequence: 1 MMASKATPSADGATGAGQ.....YOLKEVGTAGPARGRLGVR 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/US06\_PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2100	72.6	530	14	US-10-314-739-3
2	268	9.3	668	14	Sequence 3, Appli
3	264.5	9.1	547	12	Sequence 4, Appli
4	264.5	9.1	671	12	Sequence 2, Appli
5	263.5	9.1	669	14	Sequence 2, Appli
6	143.5	5.0	2971	14	Sequence 50, Appl
7	130	4.5	3930	12	Sequence 46817, A
8	126.5	4.4	1185	12	Sequence 46893, A
9	125.5	4.3	2358	12	Sequence 45763, A
10	123.5	4.3	6310	12	Sequence 67793, A
11	119.5	4.1	1231	16	Sequence 3, Appli
12	117.5	4.1	609	16	Sequence 43, Appl
13	117	4.0	1564	15	Sequence 1254, Ap
14	114.5	4.0	913	15	Sequence 2294, Ap
15	114.5	4.0	1212	16	Sequence 5, Appli

16	113.5	3.9	618	14	US-10-061-201-7
17	113.5	3.9	729	14	US-10-061-201-3
18	113.5	3.9	1147	15	US-10-327-481A-38
19	113	3.9	709	12	US-10-231-956A-287
20	113	3.9	910	16	US-10-389-566-2431
21	113	3.9	1794	10	US-09-985-728-299
22	113	3.9	1799	10	US-09-985-728-149
23	113	3.9	1821	10	US-09-965-738-82
24	113	3.9	2234	12	US-10-612-090-20
25	113	3.9	5877	14	US-10-142-515-11
26	113	3.9	5935	14	US-10-243-243A-8
27	113	3.9	11721	10	US-09-965-738-162
28	112	3.9	1023	12	US-10-147-493-200
29	112	3.9	1023	12	US-10-145-127-200
30	112	3.9	1023	12	US-10-160-503-200
31	112	3.9	1023	12	US-10-143-118-200
32	112	3.9	1023	12	US-10-144-993-200
33	112	3.9	1023	12	US-10-158-787-200
34	112	3.9	1023	12	US-10-140-024-200
35	112	3.9	1023	12	US-10-140-808-200
36	112	3.9	1023	12	US-10-152-405-200
37	112	3.9	1023	12	US-10-137-852A-200
38	112	3.9	1023	12	US-10-127-900A-200
39	112	3.9	1023	12	US-10-128-885A-200
40	112	3.9	1023	12	US-10-131-820A-200
41	112	3.9	1023	12	US-10-142-886-200
42	112	3.9	1023	12	US-10-146-728-200
43	112	3.9	1023	12	US-10-146-786-200
44	112	3.9	1023	12	US-10-147-499-200
45	112	3.9	1023	12	US-10-157-798-200

ALIGNMENTS

RESULT 1

US-10-314-739-3  
; Sequence 3, Application US/10314739  
; Publication No. US20030129588A1  
; GENERAL INFORMATION:  
; APPLICANT: ~~Estes, Mary K~~  
; Jiang, Xi  
; Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; Characterize No. US20030129588A1walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski L.L.P.  
; CITY: Washington, D.C.  
; STATE: <Unknown>  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/10/314,739  
; APPLICATION NUMBER: US/10/314,739  
; FILING DATE: 09-Dec-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311,023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX: <Unknown>

```

; INFORMATION FOR SEQ ID NO: 3:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH: 530 amino acids
;     TYPE: amino acid
;     TOPOLOGY: linear
;   MOLECULE TYPE: protein
;   SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match          72.6%; Score 2100; DB 14; Length 530;
Best Local Similarity 72.7%; Pred. No. 7.2e-182;
Matches 396; Conservative 54; Mismatches 79; Indels 16; Gaps 8;

QY 1 MMWASKDAPSPADGATGAGQLVPEVNTADPIPIDPVAGSSTALATAGQVNLIDPWIIINF 60
DB 1 MMWASKDAPSSVDGAGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINF 60
QY 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGMVRVVLGNAGFTA 120
DB 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGMVRVVLGNAGFTA 120
QY 121 GKVIICVPPGFSRSLTAQATLFPVVIADVRTLDPEVPLEDVNRVLYHNND--TQPTM 179
DB 121 GKVIICVPPGFSRSLTAQATLFPVVIADVRTLDPEVPLEDVNRVLYHNND--TQPTM 180
QY 180 RLKLCMLYPLRTGCGSGGTDSPVAVGRVLTCPGDFNFLEVPPTVEQKTRPFTVPNPL 239
DB 181 RLKLCMLYPLRTGCGGTDSPVAVGRVLTCPGDFNFLEVPPTVEQKTRPFTVPNPL 238
QY 240 KYLSNSRIPNPIGMSLSPPQTQVQFONGRCHTIDGQPLGTTPVSVLSQCKRGRITSGQ 299
DB 239 SLSNSRAPLPISSMGISPDNVQSVQFONGRCHTIDGRLVGTTPVLSHVAKIRG--TSNG 296
QY 300 RVLNLTDELGSPFMAAPAGPAPGPDGLSGCDHMSKIPNSSTQNNPIVTVNSVKPNSQQ 359
DB 297 TVINLTDELGTPFPHPFGCPAPIGFDPDLGGCDHINMTQFGHSSQTYQVDVDTT---PDT-- 351
QY 360 FVPHLSLITLDENVSSGGDYIGTIQWTSPPSDSGGANTNFWKIPDYGSSLAELASQAPAV 419
DB 352 FVPHLSGI--QANGIGSGNTVGVLSWISPPSHSPSGQVLDLWKIPNYGSSITEATHLAPSV 409
QY 420 YPPGFNEVIVFMASIPGNQSGSPNLVPCLLPOEYITHISEQAPIQGEAALLHYVDPD 479
DB 410 YPPGFGEVIVFPMKMPG---GAYNL-PCLLPOEYISLASEQAPTVGEAALLHYVDPD 465
QY 480 TNRNLGEPKLYPGGYLTCVPSNSTGPPQLDGVFVAFSWGRFVYQKVPVGTAGPARGR 539
DB 466 TGRNLGEPKAYPDGFLTCVPSNGASSGPPQLPVGIVFVSWVRFYQKVPVGTASSARGR 525
QY 540 LGVRR 544
DB 526 LGLRR 530

RESULT 2
US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-4

Query Match          9.3%; Score 268; DB 14; Length 668;
Best Local Similarity 27.6%; Pred. No. 4.3e-15;
Matches 103; Conservative 54; Mismatches 140; Indels 76; Gaps 18;

QY 71 SPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGMVRVVLGNAGFTAGKVIICVPP 130
DB 179 STSETGKILFKQSLGSLNPNVLTSLAKLYAVWAGSIEVRFISGSGVFGGKLAIVVPP 238
QY 131 GFOS-RTLSIAQATLFPVVIADVRTLDPEVPLEDVNRVLYHNNDTQPTMRLLCMLYTPL 189
DB 239 GIDPVQSTMLQ---YPHVLFDARQVEPVIFITPDLRNSLYHLMSDPTTSLVIMYNDL 295
QY 190 RTGASGTDSPVAVGRVLTCPGDFNFLEVPPTVEQKTRPFTVPNIPDKYLSNSRIPN 249
DB 296 INFYANDSNSGCIIV-TVETKPGDPFKHLLKPPG-----SMLTHGSIPS 339
QY 250 PIRGMSLS-----PDQTQNV-----QFQ-NGRCTIDGQPLG-TTPSVSVLSQCKRGR 295
DB 340 DLIPKSSSLWIGNRHMSDITDFVKPFVQANRHFDFNOETAGWSTP-----REF-PI 391
QY 296 TSGORVLNLTDELGSPFMAAP-----APAGFPDLGSCDWHIEMSKIPNSST-QNNPI 348
DB 392 T-----ITVSEKGGSKLIGVATDSIVPGIPDCMPD-----TTIPEKLTAPAGDYA 436
QY 349 VTVNSVKPNSQQFVPHLSITLDENVSSGGDYI-GTIQ--WTSPPSDSGGANTNFWKIPDY 405
DB 437 ITNGGNNDITTAADYDGASIIKNNTNPKGYICGALQRAW-----GDKKISNTAF-----I 487
QY 406 GSSLAELASQAPAV 418
DB 488 TTAIREGNSIKPS 500

RESULT 3
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-4

Query Match          9.1%; Score 264.5; DB 12; Length 547;
Best Local Similarity 33.3%; Pred. No. 6.6e-15;
Matches 73; Conservative 32; Mismatches 89; Indels 25; Gaps 7;

QY 13 DGATGAGQLVPEVNT-ADPIPIDPVAGSSTALATAGQVNLIDPWIIINFVQAPQGEFTI- 70
DB 3 DGSITA-----PEQGTWVGVIAPESQAMSTAAADMATKGSVDSEW-----EAFPSFH 49
QY 71 -----SPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGMVRVVLGNAGFTAGKVI 125
```

Db 50 TSVNWTSTQKILFKQSLGPELLLPYLEHLAKLYVAMSGSIEVRFSGSGVFGKLA 109  
Qy 126 CCVPPGFQS-RTLSIAQATLFPFHVIADVRLDPEVPELVEDVRNVLVHNDTQPTMRLLCM 184  
Db 110 IVVPPGVDPVQSTMLQ---YPHVLFDAQVEPVIFCLPDLASTLYHLMSDITDTTSLVIM 166  
Qy 185 LYTPLRTGASGSGTDSFVAGRVLCVCPGDFNFLELVPP 223  
Db 167 VYNDLINPYANDANSGCIV-TVETKPGDPDFKFLKPP 204

## RESULT 4

US-10-670-695-2  
; Sequence 2, Application US/10670695  
; Publication No. US20040058316A1  
; GENERAL INFORMATION:  
; APPLICANT: Jensen, Wayne A.  
; APPLICANT: Lappin, Michael R.  
; APPLICANT: Rosen, David K.  
; APPLICANT: Andrews, Janet S.  
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE  
; FILE REFERENCE: DI-9-1  
; CURRENT APPLICATION NUMBER: US/10/670,695  
; CURRENT FILING DATE: 2003-09-25  
; PRIOR APPLICATION NUMBER: 09/521,738  
; PRIOR FILING DATE: 2000-03-09  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 671  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-10-670-695-2

Query Match 9.1%; Score 264.5; DB 12; Length 671;  
Best Local Similarity 33.3%; Pred. No. 9.1e-15;  
Matches 73; Conservative 32; Mismatches 89; Indels 25; Gaps 7;  
Qy 13 DGATGAGQLVPEVNT-ADPIPDPVAGSSTALATAGQVNLIDPWIINNFPVQAPQGEFTI- 70  
Db 127 DGSITA----PQGTWGVGVIAEPSAQMSAADMATGKSDSEW-----EAFSFP 173  
Qy 71 ----SPNTPGDVLDLQGLPHLPFLSHLSQMYNGVGNMRVRVVLGNAFTAGKVII 125  
Db 174 TSVNWTSTQKILFKQSLGPELLLPYLEHLAKLYVAMSGSIEVRFSGSGVFGKLA 233  
Qy 126 CCVPPGFQS-RTLSIAQATLFPFHVIADVRLDPEVPELVEDVRNVLVHNDTQPTMRLLCM 184  
Db 234 IVVPPGVDPVQSTMLQ---YPHVLFDAQVEPVIFCLPDLASTLYHLMSDITDTTSLVIM 290  
Qy 185 LYTPLRTGASGSGTDSFVAGRVLCVCPGDFNFLELVPP 223  
Db 291 VYNDLINPYANDANSGCIV-TVETKPGDPDFKFLKPP 328

## RESULT 5

US-10-209-507-2  
; Sequence 2, Application US/10209507  
; Publication No. US20030109033A1  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.2  
; CURRENT APPLICATION NUMBER: US/10/209,507  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: 09/617,594  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11

; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 669  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-10-209-507-2

Query Match 9.1%; Score 263.5; DB 14; Length 669;  
Best Local Similarity 32.7%; Pred. No. 1.1e-14;  
Matches 72; Conservative 33; Mismatches 92; Indels 23; Gaps 7;  
Qy 12 ADGATGAGQLVPEVNT-ADPIPDPVAGSSTALATAGQVNLIDPWIINNFPVQAPQGEFTI 70  
Db 125 ADGSSI--TTPEQGTLVGGVIAEPSAQMAATAADATGKSDSEW-----ESFFSF 173  
Qy 71 ----SPNTPGDVLDLQGLPHLPFLSHLSQMYNGVGNMRVRVVLGNAFTAGKVI 124  
Db 174 HTSVNWTSTQKILFKQSLGPELLLPYLEHLAKLYVAMSGSVDVRFSGSGVFGKLA 233  
Qy 125 ICCVPPGFQS-RTLSIAQATLFPFHVIADVRLDPEVPELVEDVRNVLVHNDTQPTMRLLC 183  
Db 234 AIVPPGVDPVQSTMLQ---YPHVLFDAQVEPVIFSIPLRLSTLYHLMSDITDTTSLVI 290  
Qy 184 MLYTPLRTGASGSGTDSFVAGRVLCVCPGDFNFLELVPP 223  
Db 291 VYNDLINPYANDANSGCIV-TVETKPGDPDFKFLKPP 329

## RESULT 6

US-10-146-473-50  
; Sequence 50, Application US/10146473  
; Publication No. US20030108888A1  
; GENERAL INFORMATION:  
; APPLICANT: Scanlan, Matthew  
; APPLICANT: Gout, Ivan  
; APPLICANT: Stockert, Elisabeth  
; APPLICANT: Gure, Ali  
; APPLICANT: Chen, Yao-Tseng  
; APPLICANT: Old, Lloyd  
; TITLE OF INVENTION: Breast Cancer Antigens  
; FILE REFERENCE: L00461/70130(JRV)  
; CURRENT APPLICATION NUMBER: US/10/146,473  
; CURRENT FILING DATE: 2002-05-15  
; PRIOR APPLICATION NUMBER: US 60/291,150  
; PRIOR FILING DATE: 2001-05-15  
; NUMBER OF SEQ ID NOS: 82  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 50  
; LENGTH: 2971  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-146-473-50

Query Match 5.0%; Score 143.5; DB 14; Length 2971;  
Best Local Similarity 22.1%; Pred. No. 0.009;  
Matches 124; Conservative 51; Mismatches 226; Indels 159; Gaps 22;  
Qy 19 GOLVPEVNT----ADPIPDPVAGS--TALATAGQVNLIDPWIINNFPVQAPQGEFTISP 72  
Db 809 GRTVVVNNRAPLGPVDPVPPGPPELSAQPTPGVPQVLP--ASLMVVSASGAPPLIPA 866  
Qy 73 NNTPGDVLFDLQGLPHLPFLSHLSQMYNGVGNMR-----VRVVL 113  
Db 867 SRPPGPVLL-----PPLQPNSGSLPQVLPSPGLVLSGTSRPFPTTSLKPTTAPVRLSP 921  
Qy 114 AGNAFTAGKVIICCVPPGF-----QSRTLSIAQATLFPFHVIADVRLDPEVPELVED 164  
Db 922 APPGSSLLKPLTVPPGYTFPPAAATTTTATATTAVP-----APTAPQRLILSP 976  
Qy 165 VRNVLVHNDTQPTMRLLCMLYTLPLRTGASGSGTDSFVAGRVLCVCPGDFNFLELV 220

Db 977 DQARLPSGEVSVISQGLASLAQRPVANGGSKPL-TTQIQGNKLTITGAQVROLAVGQPR 1035  
Qy 221 ---VPPT-----VEQTRPFTVPNTPL-----KYLNSRIP 248  
Db 1036 PLOMPTVNNVTGVVKKIVVRQAPRDGLTFVPLAPAPRPPSSGLPAVLNPRPTLTGRLP 1095  
Qy 249 NPIEGMSLSDDQNTQVQNGRECTIDGOLGTTTPVSUSQLCKFRGRTTSQORVLNLTELD 308  
Db 1096 TPTLGTARAP-----MP-----TPTLVRPLK-----LV 1119  
Qy 309 GSPFMAFAAPAPAGPDLGSCDWHIEMS-----KIPNSQNNPNTVNSVKPNSQQ 359  
Db 1120 HSPFVSASACAPALPTISSPLHVPSSLPGPASSPMPINSSPLASPVSTVSVPLSSS 1179  
Qy 360 F-----VPLHSITLDENVSSGGDYIGTIQWTS-----PSDGGANTFWK 401  
Db 1180 LPISVPTTLPAPASAPLTTPISAPLTVSASGALLT-SVTPPLAPVVPAAAGPPLQ--- 1235  
Qy 402 IPDYGSSLAELASQAPAVP-----PGFNEVIVFMAISIPGPNOSGPNLVPCLLPOE 454  
Db 1236 -PSGAPSAALTLGLATAPSLSSSQTGPHLLAPTSSHVPGLNSTVAPACSPVLVPAS 1294  
Qy 455 YI-THFISEQAPIQGAALL 473  
Db 1295 ALASPPPSAPNPAPAQASLL 1314

RESULT 7  
US-10-282-122A-46817  
Sequence 46817, Application US/10282122A  
Publication No. US20040029129A1  
GENERAL INFORMATION:  
APPLICANT: Wang, Liangsu  
APPLICANT: Zamudio, Carlos  
APPLICANT: Malone, Cheryl  
APPLICANT: Hasebeck, Robert  
APPLICANT: Ohlsen, Kari  
APPLICANT: Zyskind, Judith  
APPLICANT: Wall, Daniel  
APPLICANT: Trawick, John  
APPLICANT: Carr, Grant  
APPLICANT: Yamamoto, Robert  
APPLICANT: Forsyth, R.  
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
FILE REFERENCE: ELITRA.034A  
CURRENT APPLICATION NUMBER: US/10/282,122A  
PRIOR FILING DATE: 2003-02-20  
PRIOR APPLICATION NUMBER: 60/191,078  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: 60/206,848  
PRIOR FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/207,727  
PRIOR FILING DATE: 2000-05-26  
PRIOR APPLICATION NUMBER: 60/230,335  
PRIOR FILING DATE: 2000-09-06  
PRIOR APPLICATION NUMBER: 60/230,347  
PRIOR FILING DATE: 2000-09-09  
PRIOR APPLICATION NUMBER: 60/242,578  
PRIOR FILING DATE: 2000-10-23  
PRIOR APPLICATION NUMBER: 60/253,625  
PRIOR FILING DATE: 2000-11-27  
PRIOR APPLICATION NUMBER: 60/257,931  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: 60/267,636  
PRIOR FILING DATE: 2001-02-09  
PRIOR APPLICATION NUMBER: 60/269,308  
PRIOR FILING DATE: 2001-02-16  
Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 78614  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 46817

LENGTH: 3930  
TYPE: PRT  
ORGANISM: Bacillus anthracis  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (19)..(19)  
OTHER INFORMATION: X=any amino acid  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (38)..(38)  
OTHER INFORMATION: X=any amino acid  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (3455)..(3455)  
OTHER INFORMATION: X=any amino acid  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (3459)..(3459)  
OTHER INFORMATION: X=any amino acid  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (3552)..(3552)  
OTHER INFORMATION: X=any amino acid  
FEATURE:  
NAME/KEY: MISC FEATURE  
LOCATION: (3566)..(3566)  
OTHER INFORMATION: X=any amino acid  
US-10-282-122A-46817

Query Match 4.5%; Score 130; DB 12; Length 3930;  
Best Local Similarity 21.1%; Pred. No. 0.23;  
Matches 136; Conservative 74; Mismatches 217; Indels 218; Gaps 33;  
Qy 4 ASKDATPSADGATGA-QGLVPEVNTA-----DPIDPVGASSTALATAGQVNLID 53  
Db 1506 ASVIATKTASSAFAAAGDTTITTLTNSGNTTANTFVFDILPAELSFVPSVQINTI- 1564  
Qy 54 PWIINNFOAPQ-----CEFTISPNTPGDVLFDLQGLPHLNPFLSHL 96  
Db 1565 -----POLGFRDPTGVPLDSIPVGGTIIISQAIVGSI-----PAINPTLNQS 1607  
Qy 97 SOMYNGWVGNMR--VRVVLGNAFTAGKVIICCVPGFQSRRTLSIAQAT-----LPHVI 149  
Db 1608 STTYSIIIVDTPQPPVTEIATSN-----PTLIQINEAITQATKSVDRLFSDV- 1653  
Qy 150 ADVRTLDPVEPLDEDVNLVYHN-NDTQPTMRLLCMLYPLRTGG-----ASGSTD 200  
Db 1654 APGNSFLTYYLLENIGNTTATNIIFTDPIPNHTVFIEDSVRVGILLPGVNPANGIPG 1713  
Qy 201 FVWAG-----RVLTCPGPDF-----NFLVLPPTVEQTRPFTV 234  
Db 1714 DIIAGDFINITFRVQVVSIPNPIFTIGPGPNSPVWNGASINYQFM-----TG 1761  
Qy 235 PNIPLKLSNRPINPIEGMSLSDDQNTQVQNGRECTIDGOLGTTTPVSUSQLCKFR- 292  
Db 1762 PNLPL--ASRSTTSNFV-----STQINSGEIAL-VKSDVKTFVTIIGDTLSSYIS 1807  
Qy 293 -----GRITSGQVRLNLTDLGSPFMAFAAPAGPDLGSCDWHIEMSKI PNSTQN--- 345  
Db 1808 LSNPGNVTSONIIFTDVLPEGITFISGTLTNDSGTQQIGNPATGIQIGNPGSTATTVI 1867  
Qy 346 NPIVTN--SVKP-----NSQQFV-----PHLSITL-----DENV 373  
Db 1868 NALVTNIPSNIPSNFSSVQFAHVVDPSQPSVSTNLNTSVTTTKSAILTTTKSGADKSV 1927  
Qy 374 SSGDYIGTIQWTSPPSDSG---GANTNFWKIPDYGSSLAELASQAPAVPFGFNEVIV 430  
Db 1928 ISVGD---TITYTTTITNTGNTAAANIKF-----TSAIPA-----NTTIFP 1965  
Qy 431 FMASIPGPNOSG-SPNL---VPCLLPQEVIT-----HFISEQAPIQGEAALLHVDPD 479  
Db 1966 NSVTINGVQSGVQPALGVNIPNIAPGETVTVTFQVNLVSPSSSSIMNDTILY----- 2020



QY 480 TNRNLGEFKLYPGGYLTCVPNSSTGPGQOLP-LDGVFVFASWVR 523  
Db 2021 -----SYTDPENG--TPVTSTSTNIVTPVLDALITMVKSDQ 2057

RESULT 8

US-10-282-122A-46893  
; Sequence 46893, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 46893  
; LENGTH: 1185  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (115)..(115)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (143)..(143)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (151)..(151)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (254)..(254)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (317)..(317)  
; OTHER INFORMATION: X=any amino acid

; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (369)..(369)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (466)..(466)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (470)..(470)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (619)..(619)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (750)..(750)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (808)..(808)  
; OTHER INFORMATION: X=any amino acid  
; US-10-282-122A-46893  
  
Query Match 4.4%; Score 126.5; DB 12; Length 1185;  
Best Local Similarity 20.9%; Pred. No. 0.077;  
Matches 126; Conservative 77; Mismatches 223; Indels 177; Gaps 30;  
  
QY 22 VPEVNT--ADPIP--IDPVAGS-----STALATAGOVNLIDPWIINNFVQAPQGEFTISPN 73  
Db 177 IPANNVIFSDPIPSWTQFVAGSVIVDGTPLPSAS-----ITSIGIGINTIEN 223  
QY 74 NTPGDVLEDLQL---GPHNLNPLSLHLSOM---YNCWGVGNMVRVVLGNAGNAFTAGKVIICC 127  
Db 224 QTV-TIIFQVQIVSNPPTFTPELQNLAFVNFXYN-----VGNALQA----- 263  
QY 128 VPPGFQSRTLSTAQATLFPPIADVRTLDPVEVPLED--VRNVLYHNNDTQPTMRLLCML 185  
Db 264 -QPG--NVETNVFVTAIHSAILSAVKTASTAFANIGDITVTVLILQNRHFTGLGX--- 317  
QY 186 YTPLRTGASGCT---DSFVVAGRVLTCPGPDFN----- 216  
Db 318 ---ISQGIPIGCTTEVENSFAVNGN--TIPGANPNSGVNIPTVSAGSSLTVTFLSRXLHS 372  
QY 217 -----FLFLVPPPTVEQKTRPFTVPNIPKXVLSNRPNPIEG----- 253  
Db 373 TPSPNPITNVASIQFAFIVDPAAPPVTCGT-T-SNSASTQINNATVTTLEADRTIVSIGD 431  
QY 254 -MSLSPDOTQNVQF--QNGRCRTIDGQPLGTPVVSQOLCKFRGRTSGORVNLNLTLDGSP 311  
Db 432 IITYTATLTNTGNFPANSVLLINGVPEGALFVPIX--CHAXRDTFSEHQV-----QLSGIP 485  
QY 312 -----FMAFAAPAPAGP---PDLGSCDWHIEMSKIIPNSSTQNPNIVTNSV 353  
Db 486 VGI IAPGNSATITTFQFLANSIP-PQGAIIINQALTSYTVIVDPSQPPVATSSNTVTTAV 544  
QY 354 KPNQQFVPHLSSITLDENVSSGGDIYIGTIQWTSPPSDSGGANTFWKIPDYGVSSLAELAS 413  
Db 545 VDASLSVKNITDSIVQSTD-----GTITYTVVVIQNNGTANTVTLTDL---VPEGT 593  
QY 414 OLAPAVYPPGFNEVIVYFMASIPGPNQSGSNLVPCLLPQEVITHFISEQAPIQGEAALL 473  
Db 594 ALIP-----NSVTINSI-SIPGGRSKGKSHXTSLAPSEIVPGHIPKLSILLFQSG-- 643  
QY 474 HYVDPDTNRNLGEPKLYPGGYLTCVPNSSTGPGQOLP-LDGVFVFASWVR 533  
Db 644 -----NPIFNARY----GLYFYCGSNCSYCLEQL-----LRNQLSHKFMQRMVFL 687  
QY 534 GPA 536  
Db 688 KPS 690

Db 366 VPTGTGSAENVVLRDSIPN-----GTFVAGSVTVGGVTPQNPANPATGINL 412  
QY 279 GTTPVSVSQLCKFRGRTSGQRVLNLTLDGSPFMAFAPAPAGFPDLGSCDWHIEMSKI 338  
Db 413 GTIPNNTQIRIVTFQVRITS-----FNPENP-----I 438  
QY 339 PMSSTQN-----NPVTSVKPNSQOFVPHLSSITLDENVVS-SGGDYIGTIGTQWTSPP 389  
Db 439 PNRAMVSYQFRPRFVGSPLITSSSSNTVQTTVQATISMQKSVDLQATLNDVLTVTNVV 498  
QY 390 SDSCGANTNFKIPDYGSSLAELASQAPAVYPPGNEVIVYFMASIPGPNQSGSNLYPC 449  
Db 499 TNGNVTAN--NVIFVDSIPAGTTFVA-----NSVIVNGVAR-PGANPASSINLGI 547  
QY 450 LLPQEVITHIFISEQAPIQGEAALLHVDPDTHNRLGEEKLYPGGYLTCVPSNSSSTGPOOL 509  
Db 548 NASQTTVVR-----QVRVTSNPLVNPINRASATFTFP-----VPGQQ 587  
QY 510 PLDG 513  
Db 588 PVSG 591

RESULT 10  
US-10-282-122A-67793  
; Sequence 67793, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 67793  
; LENGTH: 6310  
; TYPE: PRT  
; ORGANISM: Pseudomonas putida  
US-10-282-122A-67793

US-10-282-122A-45763  
; Sequence 45763, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 45763  
; LENGTH: 2358  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
US-10-282-122A-45763

Query Match 4.3%; Score 125.5; DB 12; Length 2358;  
Best Local Similarity 22.1%; Pred. No. 0.27;  
Matches 120; Conservative 57; Mismatches 174; Indels 193; Gaps 28;

QY 18 ACQVPEVNTADPID-PVAGSSTAL--ATAGVNLDPMLNMFVQAPQEGFTISPN 74  
Db 193 AGQ---EIIDASSPVDATISGFATPATGAVTGRV-----LVTAQSGDSNIS--- 236  
QY 75 TPGDVLFDLQGLPHLNPFLS-----HLSQWNGWGNMVRVVLGNAPTAGKV 123  
Db 237 --GD---QURFGENATTVALLGPRNPNANFFQSIQICND-----SGNLDTSQTF 280  
QY 124 IICCVPPGQSRTLSIAQNTLPHVADVRLDPEVPLEVVRNV-----LYHNNDTQPTM 179  
Db 281 -----GDLNQLPLGVALA-----VARQGWDTNVDASSLVNNTSIV 318  
QY 180 RULICMLYTLRTGASGGTDSFVVA--GRVLTCPGDFNLFLEVPPTV-----EQKTRPFT 233  
Db 319 RFV-----TNG-----DGYAAAGFGVQIDATGPIINPVKSVNRTVAGVGDTLTYTIT 365  
QY 234 VFNIPILKLSN-----SRIPNPIEGMSLSPDQTONVQFQNGRCTIDG--QP-----L 278



Wed Jun 2 09:13:28 2004

```

; PRIOR APPLICATION NUMBER: 60/243,863
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/244,443
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 60/245,029
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,995
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/245,293
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/245,315
; PRIOR FILING DATE: 2000-11-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 43
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Xenopus laevis
; ORGANISM: Xenopus laevis
US-10-055-569A-43

Query Match 4.1%; Score 117.5; DB 16; Length 609;
Best Local Similarity 20.8%; Pred. No. 0.18; Indels 165; Gaps 24;
Matches 96; Conservative 44; Mismatches 156;

Qy 23 PEVNTADP-----IPIDPVAGSSTALATAGQVNLIDPW-----IINNFOAPQGEFTISP 72
Db 265 PVAFADPWGASAAAPADPWAGGATPASVFAAAAADPPWGGPPVATGSSSDPWGT-GVQT 323
Qy 73 NTPPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAGTACKVILCCVP--P 130
Db 324 NSTPGD-----PWGDT-----QAVTSADKVSVDPNP 351
Qy 131 GFQSRITLSIAQA--TLPPHVIADV-RTLDPVEVEPLEDVRNVLHNNDTQPTMRLLCMLYT 187
Db 352 GGSATTATPSDDPWSPPVAQSVKKAADPWAPAAFSFSDP--WGGSPSKP-----400
Qy 188 PLRTGGASGGTDSFVAVAGRVLTGCGDFNLFVLPVPTVEQKTRPFTVNIPLKYLNSRI 247
Db 401 --NTNGTMGELD--LLAGEV-----PMSRSLGSKS 426
Qy 248 NPTEGMSLSDQTNQVQFQNGRCTIDGQPLGTTFVSVSQLCCKPRGRTISGORVLNLT 307
Db 427 PDADFDMSTMS-----GSLCDF-----NPTRK 448
Qy 308 DGSFPMFAAPAPAPAGFPDLGSCDWHIEMSKIPNSSTON-----NPITNSVKNSQ 359
Db 449 TPESFLG---PNAALVDLSL---ISKSTLQNTKTSNPFVLTGTPTATNPFPQPNQ--500
Qy 360 FVPHLSSTLDENVSSGGDYGTGTQWTSPPSDSGANTFWKIPDYGSSLAELASQAPAV 419
Db 501 -----SSLTNQLRSPVMTLG-----QVTPAGQTAT---IP-PASPMSVSMPAGI 546
Qy 420 YPPGFNEVI-VYFMAISIP-GPNQSGSNLV--PCILLPOEYI 456
Db 547 PLANMAPVGMQPMAGVPVGTILAPGVGWLPPMPPQOLV 587

RESULT 13
US-10-292-798-1254
; Sequence 1254, Application US/10292798
; Publication No. US20030235833A1
; GENERAL INFORMATION:
; APPLICANT: SUWA, MAKIKO
; APPLICANT: ASAI, KIYOSHI
; APPLICANT: AKIYAMA, YUTAKA
; APPLICANT: ABURATANI, HIROYUKI
; TITLE OF INVENTION: GUANOSINE TRIPHOSPHATE-BINDING PROTEIN COUPLED RECEPTORS
; FILE REFERENCE: 084335/166
; CURRENT APPLICATION NUMBER: US/10/292,798
; PRIOR FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: 10/017,161
; PRIOR FILING DATE: 2001-12-18

; PRIOR APPLICATION NUMBER: 60/243,863
; PRIOR FILING DATE: 2000-10-27
; PRIOR APPLICATION NUMBER: 60/244,443
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 60/245,029
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/244,995
; PRIOR FILING DATE: 2000-11-01
; PRIOR APPLICATION NUMBER: 60/245,293
; PRIOR FILING DATE: 2000-11-02
; PRIOR APPLICATION NUMBER: 60/245,315
; PRIOR FILING DATE: 2000-11-02
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 137
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 43
; LENGTH: 609
; TYPE: PRT
; ORGANISM: Xenopus laevis
; ORGANISM: Xenopus laevis
US-10-055-569A-43

Query Match 4.1%; Score 117.5; DB 16; Length 609;
Best Local Similarity 20.8%; Pred. No. 0.18; Indels 165; Gaps 24;
Matches 96; Conservative 44; Mismatches 156; Indels 165; Gaps 24;

Qy 23 PEVNTADP-----IPIDPVAGSSTALATAGQVNLIDPW-----IINNFOAPQGEFTISP 72
Db 265 PVAFADPWGASAAAPADPWAGGATPASVFAAAAADPPWGGPPVATGSSSDPWGT-GVQT 323
Qy 73 NTPPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAGTACKVILCCVP--P 130
Db 324 NSTPGD-----PWGDT-----QAVTSADKVSVDPNP 351
Qy 131 GFQSRITLSIAQA--TLPPHVIADV-RTLDPVEVEPLEDVRNVLHNNDTQPTMRLLCMLYT 187
Db 352 GGSATTATPSDDPWSPPVAQSVKKAADPWAPAAFSFSDP--WGGSPSKP-----400
Qy 188 PLRTGGASGGTDSFVAVAGRVLTGCGDFNLFVLPVPTVEQKTRPFTVNIPLKYLNSRI 247
Db 401 --NTNGTMGELD--LLAGEV-----PMSRSLGSKS 426
Qy 248 NPTEGMSLSDQTNQVQFQNGRCTIDGQPLGTTFVSVSQLCCKPRGRTISGORVLNLT 307
Db 427 PDADFDMSTMS-----GSLCDF-----NPTRK 448
Qy 308 DGSFPMFAAPAPAPAGFPDLGSCDWHIEMSKIPNSSTON-----NPITNSVKNSQ 359
Db 449 TPESFLG---PNAALVDLSL---ISKSTLQNTKTSNPFVLTGTPTATNPFPQPNQ--500
Qy 360 FVPHLSSTLDENVSSGGDYGTGTQWTSPPSDSGANTFWKIPDYGSSLAELASQAPAV 419
Db 501 -----SSLTNQLRSPVMTLG-----QVTPAGQTAT---IP-PASPMSVSMPAGI 546
Qy 420 YPPGFNEVI-VYFMAISIP-GPNQSGSNLV--PCILLPOEYI 456
Db 547 PLANMAPVGMQPMAGVPVGTILAPGVGWLPPMPPQOLV 587

RESULT 14
US-10-369-493-2294
; Sequence 2294, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
```

```

; PRIOR APPLICATION NUMBER: JP 2001-246789
; PRIOR FILING DATE: 2001-06-18
; NUMBER OF SEQ ID NOS: 2070
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 1254
; LENGTH: 1564
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-10-292-798-1254

Query Match 4.0%; Score 117; DB 15; Length 1564;
Best Local Similarity 20.8%; Pred. No. 0.87;
Matches 115; Conservative 50; Mismatches 181; Indels 208; Gaps 27;

Qy 10 PSADGATGAGQVPEVNTADPIPIDPVAGSSTALATAGQVNLIDPWVQAPQGEFT 69
Db 1160 PHARASAGICSLVPS-----VCRSPASLAVAALGSGV-----CHGGAFS 1198
Qy 70 ISPNNTPGDVLFDLQ---LQPHLNPFLSHLSQMYNGWGNMVRVVLGNAGTACKVILC 126
Db 1199 CS-----LVDCQGRGDGAGARELLSLPQGGSGYGN-----LVP 1233
Qy 127 CVPPGFSQRTLSIAQATLPPHVIADYRTLDPVEVEPLEDVRNVLHNNDTQPTMRLLCMLY 186
Db 1234 QSPPPVPSRTVPSVTSSVPPTVPSVPS--SSCAVPHSPHPLH-----LCPLH 1281
Qy 187 TELRTGASGCTDSFVAVAGRVLTCPGDFNLFVLPVPTVEQKTRPFTVNIPLKYLNSR 246
Db 1282 SPLSP-----LHLLCPPTVPSVPSVSCVLPQSP-----QS 1310
Qy 247 IPNPIEGMS-----LSPDQT-QNVQFQNGRCTIDGQPLGTTFVSVSQLCCKPRGRTISQ 301
Db 1311 SPPVPSQSSPPVPSPPQSPQSSPPVSSSTVSVPSSCVPTVPS-----VTSSCA 1363
Qy 302 LNLTELDGSPWFAAPAPAGFPDLGSCDWHIEMSKI-----PNSSTQNNP-----347
Db 1364 L-----PHSPLHHP--LLLCPPHSPHLLCCPPAQSPQSPPPVPSRTVPS 1408
Qy 348 IVTNSVKNSQFVPHLSSTLDENVSSGGDYGTGTQWTSPPSDSGANTFWKIPDYGS 407
Db 1409 VTSSSVPPQSPQSPPPSPPL-----SPPS-----PVPSPQS 1439
Qy 408 SLAELASQAPAVYPPGFNEVIYFMAISIPGNQSGS-----PNLVP-----CL 450
Db 1440 PLSPPPVSPPCI-PP-----VTSFSCASPPQSPQSPPLPPVPHPLHLLNAPPHRGAVAL 1493
Qy 451 LPQEYITHFISQ---APIQGEAALLHYVDPTN--RNIGFELYPGGYLTCTVPSN--502
Db 1494 LPAPHGASKLHQRDLLPGGRSELLQWVLPYIQHPCARGE---PG---TCCPGQGNHV 1546
Qy 503 -----STGPOQ 508
Db 1547 QWPPRAALSGLPGQ 1560

RESULT 14
US-10-369-493-2294
; Sequence 2294, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
```

SEQ ID NO 2294  
LENGTH: 913  
TYPE: PRT

ORGANISM: Schizosaccharomyces pombe

FEATURE:

NAME/KEY: unsure

LOCATION: (1)-(913)

OTHER INFORMATION: unsure at all Xaa locations

US-10-369-493-2294

Query Match 4.0%; Score 114.5; DB 15; Length 913;  
Best Local Similarity 19.8%; Pred. No. 0.64;  
Matches 117; Conservative 96; Mismatches 212; Indels 165; Gaps 32;

QY 9 TPSAGATGA--GQLVPEVNTADPIPID-----PVAGSSTA-----LATAG--QVN 50  
Db 243 TVPAVGSTTSTVPDVVXAGNVEVTEPDIMVSVTVPAVGSTTSTVPDVPXAGIVEVN 302  
QY 51 LIDPWLNNFVQAPQGEFTI--SPNNTPGDVLGLPHLNPFLS--HLSQMYNGWGNMR 108  
Db 303 GIDPETIVSVTVPAVGSTTMTAPDEVXAVIVDIK---ETDPEITVSVTVPAVGSTTSTV 359  
QY 109 VRV--LAGNAP---TAGKVIICVPPGFSRTLSIAQATLFPHVI-----ADVRLTDPVE 159  
Db 360 TRVPLAGNVEVYPTETPMVSVTVPAVGSTTSTV-----PEVPLAGNVEVKLEPET 413  
QY 160 -----VPLEDVNRNLYHNNDTQPTMRLLCMLYTPRLTGASGGTDS 200  
Db 414 MVSIVPAVGSTTSTVPEVPLAGNVEVNALEPDT-----MVSIVP-----AVGSTTS 462  
QY 201 FV-----VAGRLTCGPDNFNLFVPPVVEOKTRPFTVPNPLKYLNSRIPNPIEG-M 254  
Db 463 TVPEVPLAGNVEVYPTETPMVSVTVPAVGSTTSTV-----TVPEV-----PLAGNV 509  
QY 255 SLSPDOTQNVQFNGRCTIDGQPLGTPVSVLSOLCKFRGRI-----TSQORVNLNLTLDGS 310  
Db 510 EYVTEPETM-----VSVTVPAVGSTTSTVPEVPLAGNVEVYPTETMVSIV----- 558  
QY 311 PFMAFAAPAGAPFDLGSCDWHIEMSKIPNSSTQNNPIVTVNSVKPNSS--QQFVPHLSSTI- 367  
Db 559 -----VPAVGSTN-----STVPEVPLAGNVEVNALEPDTMVSIVPAVGSTT 601  
QY 368 -TLDENVSSGGDYIGTIQWTSPPSDGGANTFWKIPDYGSSLAESAQLAP-----AVYP 421  
Db 602 STVPEVPLAGNV--EVKLEPETMVSIV-----VPAVGSTTSTVPEVPLAGNVEVYP 653  
QY 422 PGFNEVIVYFMASIPGNOSGSPNLVP-----CLLPQEVY-----THFISEQ 463  
Db 654 TB-PETMVSIVPAVGSTTSTVPEVPLAGNVEVKLEPETMVSIVPAVGSTTSTVPEV 712  
QY 464 APIQGEAALLHYVDPDTRNRLGFEKLYPGGYLTCVPNSSSTGPOQLPLDG 513  
Db 713 VPLAGNVE-VNALEPDTMVSIV-----IVPAVGSTTSTVPEVPLAG 752

RESULT 15

US-10-618-581-5

Sequence 5, Application US/10618581

Publication No. US2004007524A1

GENERAL INFORMATION:

APPLICANT: MANNING, GERARD

APPLICANT: SUDARSANAM, SUCHA

TITLE OF INVENTION: METHOD FOR SELECTIVELY INHIBITING FUNGAL GROWTH

FILE REFERENCE: 034536/0323

CURRENT APPLICATION NUMBER: US/10/618,581

CURRENT FILING DATE: 2003-07-15

PRIOR APPLICATION NUMBER: 60/395,624

PRIOR FILING DATE: 2002-07-15

NUMBER OF SEQ ID NOS: 42

SOFTWARE: PatentIn Ver. 2.1

SEQ ID NO 5

LENGTH: 1212

TYPE: PRT

ORGANISM: Candida albicans

US-10-618-581-5

Query Match 4.0%; Score 114.5; DB 16; Length 1212;  
Best Local Similarity 20.4%; Pred. No. 0.99;  
Matches 114; Conservative 72; Mismatches 215; Indels 157; Gaps 29;

QY 23 PEVN--TADPIDPVAGSSTALATAGQVNLIDPWIINNFVQAPQGEFTISN----- 73  
Db 591 PEVSQPLPEPKVPPPEIINPAVATQAQANMTAPKIVETSETPQR--TLDSKQSVDBK 648  
QY 74 -NTPGDVLPDLQLGPHLNPFLSHLSQMYNGWGNMRVVRVVLGNAGTACKVIICCVPPGF 132  
Db 649 PSAPGP-----STAVPEQAHTTS-----VPSSE 671  
QY 133 QSRTLSIAQATL-FPHV-----IADVRLTDPVEVPLEDVNRNLYHNNDTQPTMRLLCMLYT 187  
Db 672 VKTQTSIDEDQLSIEPQQSPRTSTPTQLDPAKV----- 704  
QY 188 PLRTGASGGTDSFVAGRVLTCPGPDNFNLFVPPVTEQKTRPFTVPNPLKYLNSRI 247  
Db 705 ---VGGSGSAISAPNAG-----SGAFNSLLR-----RLSSKVKYGASSPKR--STSPS 749  
QY 248 NPPIEGMSLSPDOTQNVQFONG---RCTIDGQPLGTPVSVLSOLCKFR---GRITSGQRV 301  
Db 750 PN-VEGLSPQPTKA--DPMVRGVSVMKVTAKQKQTNTRPPK-SELIKKQPHGRSSS----- 802  
QY 302 LNLTELDGSPPFMAFAAPAGAPFDLGSCDWHIEMSKIPNSSTQNNPIVTVNSVKPNSSQQPV 361  
Db 803 -TSNKKQGFIPVEYLP-----LPTIDTNTVTVSDGAKQQLNT-V 841  
QY 362 P-----HLSSITLDENVSSG---GDYIGTIQWTS-----PPSDGGANTFWKII--PDYGS 408  
Db 842 PSTARHMHPTARAKSVGGCHMRKDSYGRVSHGSQNPPLPPLPTSMASQNSQEVVVKDTSBG 901  
QY 409 LAEASQALAPAVYP--PGFNEVIVYFMASIPGNOSGSPNLVPCLLPOEVIHFISEQAPI 466  
Db 902 PFDDVQLDDVGQYEPQLTESIIEQYNISKFN--SMPSIEHC--KTLFLKGFPSVQTTS 957  
QY 467 QGEAALLHYVDPDTRNRLG-EFKLYPGGYLTCV-----PNSSSSTGPOQLPLDGVPVFAS 519  
Db 958 AKPLPVIRYININVLKGVKQFQVKGGFV-CMHTPSVQPSVQPSHNSDELDEENKLYGDAFKSK 1016  
QY 520 WVSRYQLKPVGTAGPAR 537  
Db 1017 SSDSPEAAEPEGSKTFSR 1034

Search completed: June 1, 2004, 14:04:33

Job time : 36.7758 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.4966 Seconds  
(without alignments)  
4985.230 Million cell updates/sec

Title: US-09-926-799-4  
Perfect score: 2891  
Sequence: 1 MMASKDATPSADGATGAGQ.....YOLKPVGTAGPARGRLGVR 544

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2234.5	77.3	546	B37491	major capsid prote
2	2096	72.5	530	B37471	capsid protein - N
3	1171.5	40.5	542	S60616	capsid protein - h
4	1154	39.9	539	S40111	capsid protein - h
5	347.5	12.0	576	A53982	capsid protein - E
6	304	10.5	2344	RRWRH	genome polyprotein
7	298	10.3	2344	S53399	genome polyprotein
8	297	10.3	2344	S64740	genome polyprotein
9	285	9.9	703	C48562	coat protein - San
10	281	9.7	702	A48562	coat protein - San
11	274	9.5	668	VCWVFC	coat protein - fel
12	272.5	9.4	668	JQ2354	capsid protein - f
13	264.5	9.1	671	VCWVFF	coat protein - fel
14	263.5	9.1	668	VCWVFF	coat protein - fel
15	253	8.8	668	JQ2356	capsid protein - f
16	148	5.1	2194	GNNYE7	genome polyprotein
17	141	4.9	2508	S61441	surface-associated
18	137	4.7	2206	S03822	genome polyprotein
19	136.5	4.7	3473	A46112	genome polyprotein
20	136.5	4.7	3473	S27927	polyprotein - rice
21	135	4.7	2206	GNNY4P	genome polyprotein
22	134	4.6	2175	GNNYFE	genome polyprotein
23	132.5	4.6	2206	S3327	genome polyprotein
24	123	4.3	915	S36327	clathrin assembly
25	122	4.2	2205	GNNY2W	genome polyprotein
26	122	4.2	2207	GNNY5P	genome polyprotein
27	121	4.2	2183	T42764	coagulation factor
28	120.5	4.2	896	S63326	clathrin assembly
29	119.5	4.1	1231	S30185	insulin receptor s

30	117.5	4.1	1235	1	S16948	insulin receptor s
31	116.5	4.0	2207	2	S09553	genome polyprotein
32	116	4.0	3016	2	S77300	hypothetical prote
33	115.5	4.0	929	2	A44048	genome polyprotein
34	115.5	4.0	1367	2	T33819	hypothetical prote
35	115.5	4.0	2179	1	GNNYH4	genome polyprotein
36	113.5	3.9	745	2	G72453	hypothetical prote
37	113.5	3.9	799	2	S75351	hypothetical prote
38	113.5	3.9	2333	1	GNNY2F	genome polyprotein
39	113.5	3.9	3085	2	T00327	polyprotein - infe
40	113	3.9	1075	2	A57377	transcription fact
41	113	3.9	2150	1	GNNYH2	genome polyprotein
42	113	3.9	2206	1	GNNY21	genome polyprotein
43	112.5	3.9	537	2	T04822	hypothetical prote
44	112.5	3.9	648	2	PC4395	mucin 3 - human (f
45	112	3.9	554	2	T01240	laccase (EC 1.10.3

ALIGNMENTS

RESULT 1

B37491  
- major capsid protein [similarity] - Southampton virus  
NAlternate names: orf2 protein  
C:Species: Southampton virus  
C>Date: 22-Sep-1993 #sequence\_revision 25-Apr-1997 #text\_change 18-Aug-2000  
C:Accession: B37491  
R:Lambden, P., Ravi Gaur, J., Ashley, C.R.; Clarke, I.N.  
Sequence 259, S16948-1993  
Afflitter: Sequence and genome organization of a human small round-structured (Norwalk-like) virus  
A:Reference number: A37491; MUID:93142023; PMID:8380940  
A:Accession: B37491  
A:Status: not compared with conceptual translation  
A:Molecule type: nucleic acid  
A:Residues: 1-546 <I>AM>  
A:Cross-references: GB:L07418; NID:g1236787; PIDN:AAA92984.1; PID:g295114  
A:Note: sequence extracted from NCBI backbone (NCBIP:123458)  
A:Note: small round-structured virus, SRSV, Norwalk virus, Norwalk-like virus, serotype  
C:Superfamily: human calicivirus capsid protein  
C:Keywords: glycoprotein  
F:303,340,441/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 77.3%; Score 2234.5; DB 2; Length 546;

Best Local Similarity 74.9%; Pred. No. 1e-149;

Matches 408; Conservative 58; Mismatches 78; Indels 1; Gaps 1;

QY	1	MMASKDATPSADGATGAGOLVPEVNTADP	1	MMASKDATPSADGATGAGOLVPEVNTADP	60
DB	1	MMASKDAPQSDAGSAGOLVPEVNTADP	1	MMASKDAPQSDAGSAGOLVPEVNTADP	60
QY	61	VQAPQGEFTISPNNTPGDVLFDQLGPHLNP	61	VQAPQGEFTISPNNTPGDVLFDQLGPHLNP	120
DB	61	VQSPQGEFTISPNNTPGDVLFDQLGPHLNP	61	VQSPQGEFTISPNNTPGDVLFDQLGPHLNP	120
QY	121	GRVVICCVPPGQSRTLSIAQATLFP	121	GRVVICCVPPGQSRTLSIAQATLFP	180
DB	121	GKIIIVCCVPPGPTSSSLTIAQATLFP	121	GKIIIVCCVPPGPTSSSLTIAQATLFP	180
QY	181	LLCMLYTLRTGCGAGGTTDSFVAGRVLT	181	LLCMLYTLRTGCGAGGTTDSFVAGRVLT	240
DB	181	LVCMLYTLRTGCGGSGNSDSFVAGRVLT	181	LVCMLYTLRTGCGGSGNSDSFVAGRVLT	240
QY	241	YLSNRIINPTEGMSLSPDQTONVQNGRC	241	YLSNRIINPTEGMSLSPDQTONVQNGRC	300
DB	241	TLSNRFPSLQGMILSPDASQVQNGRC	241	TLSNRFPSLQGMILSPDASQVQNGRC	300
QY	301	VLNLTDLGSPFMFAFAPAGFPDLGSCD	301	VLNLTDLGSPFMFAFAPAGFPDLGSCD	360
DB	301	TLNLTVDGKPMFADSPAPVGFDFGKCD	301	TLNLTVDGKPMFADSPAPVGFDFGKCD	360
QY	361	VPHLSITLDENVSSGGDYTGITQNTW	361	VPHLSITLDENVSSGGDYTGITQNTW	419

361 VPHLGSIOFDEVNHTGDICTIEWISQSPSTPGTDINLWEIPDYGSSLSQAANLAPPV 420  
 420 YPGFNEVLYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479  
 421 FPGFGEALVYFVSAPFNNRNSAPNVPCLLPQYIYTHFISEQAPTGWDAALLHYVDPD 480  
 480 TNRNLGEFLYPGGYLTCPVNSSSTGPOQLPLDGVVFVFWASVSRFYQLKPVGTAGPARGR 539  
 481 TNRNLGEFLYPGGYLTCPVNGVAGPQQLPLNGVFLFVSWVSRFYQLKPVGTASTARGR 540  
 540 LGVRR 544  
 541 LGVRR 545

RESULT 2  
 B37471  
 capsid protein - Norwalk virus  
 C:Species: Norwalk virus  
 C:Date: 24-Feb-1994 #sequence\_revision 25-Apr-1997 #text\_change 28-Jul-2000  
 C:Accession: B37471  
 R:Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.  
 Virology 195, 51-61, 1993  
 A:Title: Sequence and genomic organization of Norwalk virus.  
 A:Reference number: A37471; MUID:93303939; PMID:8391187  
 A:Accession: B37471  
 A:Status: preliminary; not compared with conceptual translation  
 A:Molecule type: genomic RNA  
 A:Residues: 1-530 <JFA>  
 A:Cross-references: GB:M87661; NID:g1061311; PIDN:AA850466.1; PID:g1061313  
 A:Note: sequence extracted from NCBI backbone (NCBIP:134157)  
 C:Superfamily: human calicivirus capsid protein

Query Match 72.5%; Score 2096; DB 2; Length 530;  
 Best Local Similarity 72.5%; Pred. No. 5.6e-140;  
 Matches 395; Conservative 55; Mismatches 79; Indels 16; Gaps 8;

QY 1 MMASKDATSADGATGAGQLVPEVNTADIPIDPVAGSSALATAGQVNLDPWIINNF 60  
 DB 1 MMASKDATSADGATGAGQLVPEVNTADIPIDPVAGSSALATAGQVNLDPWIINNF 60

QY 61 VOAPQGETISPNTPGDVLFDLQGLHLPFLSHLSQMYNGWGNRVVWLAGNAFTA 120  
 DB 61 VOAPQGETISPNTPGDVLFDLQGLHLPFLSHLSQMYNGWGNRVVWLAGNAFTA 120

QY 121 GKVIICCPVPGFQSGRTLSIAQATLFPVHVIADVRLDVEVPLEVDRVNLVHND-TQPTM 179  
 DB 121 GKVIICCPVPGFQSGRTLSIAQATLFPVHVIADVRLDVEVPLEVDRVNLVHND-TQPTM 179

QY 180 RLLCMLYPLRTGASGTDTSFVAVAGVLTCPGDFNLFVLPVPTVEQKTRPFTVNIPL 239  
 DB 180 RLLCMLYPLRTGASGTDTSFVAVAGVLTCPGDFNLFVLPVPTVEQKTRPFTVNIPL 239

QY 240 KYLSNRIPNPIEGMSLSPDOTQNVQFQNGRCTLGQPLGTPVSVSOLCKFRGRIISGQ 299  
 DB 240 KYLSNRIPNPIEGMSLSPDOTQNVQFQNGRCTLGQPLGTPVSVSOLCKFRGRIISGQ 299

QY 360 FVPHLSSTILDENVSSGDYICTIOWTSPDSSGANTFNWKIPDYGSSLAESAOLAPAV 419  
 DB 360 FVPHLSSTILDENVSSGDYICTIOWTSPDSSGANTFNWKIPDYGSSLAESAOLAPAV 419

QY 420 YPGFNEVLYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479  
 DB 420 YPGFNEVLYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479

QY 480 TNRNLGEFLYPGGYLTCPVNSSSTGPOQLPLDGVVFVFWASVSRFYQLKPVGTAGPARGR 539  
 DB 480 TNRNLGEFLYPGGYLTCPVNSSSTGPOQLPLDGVVFVFWASVSRFYQLKPVGTAGPARGR 539

QY 490 TNRNLGEFLYPGGYLTCPVNSSSTGPOQLPLDGVVFVFWASVSRFYQLKPVGTAGPARGR 539  
 DB 490 TNRNLGEFLYPGGYLTCPVNSSSTGPOQLPLDGVVFVFWASVSRFYQLKPVGTAGPARGR 539

540 LGVRR 544  
 541 LGVRR 545

RESULT 3  
 S60616  
 capsid protein - human calicivirus (strain Melksham)  
 C:Species: human calicivirus  
 C:Variety: strain Melksham  
 C:Date: 23-May-1997 #sequence\_revision 23-May-1997 #text\_change 28-Jul-2000  
 C:Accession: S60616  
 R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.  
 Virus Res. 37, 271-283, 1995  
 A:Title: Capsid diversity in small round-structured viruses: molecular characterization  
 A:Reference number: S60615; MUID:96136658; PMID:8533462  
 A:Accession: S60616  
 A:Molecule type: genomic RNA  
 A:Residues: 1-542 <GRE>  
 A:Cross-references: EMBL:X81879; NID:g976077; PIDN:CAA57462.1; PID:g976079  
 A:Experimental source: strain Melksham  
 A:Note: it is uncertain whether Met-1 or Met-3 is the initiator  
 C:Superfamily: human calicivirus capsid protein  
 C:Keywords: capsid protein; coat protein

Query Match 40.5%; Score 1171.5; DB 2; Length 542;  
 Best Local Similarity 45.6%; Pred. No. 7.8e-75;  
 Matches 256; Conservative 87; Mismatches 176; Indels 43; Gaps 14;

QY 1 MMASKDATSADGATGAGQLVPEVNTADIPIDPVAGSSALATAGQVNLDPWIINNF 60  
 DB 1 MMASKDATSADGATGAGQLVPEVNTADIPIDPVAGSSALATAGQVNLDPWIINNF 60

QY 61 VOAPQGETISPNTPGDVLFDLQGLHLPFLSHLSQMYNGWGNRVVWLAGNAFTA 120  
 DB 61 VOAPQGETISPNTPGDVLFDLQGLHLPFLSHLSQMYNGWGNRVVWLAGNAFTA 120

QY 121 GKVIICCPVPGFQSGRTLSIAQATLFPVHVIADVRLDVEVPLEVDRVNLVHND-TQPTM 179  
 DB 121 GKVIICCPVPGFQSGRTLSIAQATLFPVHVIADVRLDVEVPLEVDRVNLVHND-TQPTM 179

QY 180 RLLCMLYPLRTGASGTDTSFVAVAGVLTCPGDFNLFVLPVPTVEQKTRPFTVNIPL 239  
 DB 180 RLLCMLYPLRTGASGTDTSFVAVAGVLTCPGDFNLFVLPVPTVEQKTRPFTVNIPL 239

QY 240 KYLSNRIPNPIEGMSLSPDOTQNVQFQNGRCTLGQPLGTPVSVSOLCKFRGRIISGQ 299  
 DB 240 KYLSNRIPNPIEGMSLSPDOTQNVQFQNGRCTLGQPLGTPVSVSOLCKFRGRIISGQ 299

QY 298 ---GQVNLTELDSPP-MAFAAPAPAGPDL-GSCDWHIEMSKIPINSSTONNPIVT- 350  
 DB 298 ---GQVNLTELDSPP-MAFAAPAPAGPDL-GSCDWHIEMSKIPINSSTONNPIVT- 350

QY 351 -NSVKPN-SQQVPHLSSTILDENVSSGDYICTIOWTSPDSSGANTFNWKIPDYGSSLAESAOLAPAV 419  
 DB 351 -NSVKPN-SQQVPHLSSTILDENVSSGDYICTIOWTSPDSSGANTFNWKIPDYGSSLAESAOLAPAV 419

QY 399 FWKIPDYGSSLAESAOLAPAVPPGFNEVLYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479  
 DB 399 FWKIPDYGSSLAESAOLAPAVPPGFNEVLYFMASIPGNSQSNLVPCLLPQYIYTHFISEQAPIQGEAALLHYVDPD 479

QY 404 QWVVPYAGALNLTNLPASVAPVFFGERLLFRSHLPLKGGYGNP-AIDCLLPQEWQVH 462  
 DB 404 QWVVPYAGALNLTNLPASVAPVFFGERLLFRSHLPLKGGYGNP-AIDCLLPQEWQVH 462

QY 459 FISEQAPIQGEAALLHYVDPDNRLNLPGLVPGGYLTCPVNSSSTGPOQLPLDGVVFVFA 518  
 DB 459 FISEQAPIQGEAALLHYVDPDNRLNLPGLVPGGYLTCPVNSSSTGPOQLPLDGVVFVFA 518

QY 463 FYQEAAPSSEVALVRIINPDTCRALFEAKLHRAQMTVSSNTSA--PVVVPANGVFRPD 520  
 DB 463 FYQEAAPSSEVALVRIINPDTCRALFEAKLHRAQMTVSSNTSA--PVVVPANGVFRPD 520

QY 519 SWVSRYFQKPKVGTAGPARGR 540  
 DB 519 SWVSRYFQKPKVGTAGPARGR 540

QY 521 SWVNFYSLAPMCT-GNGRRRV 541  
 DB 521 SWVNFYSLAPMCT-GNGRRRV 541

RESULT 4  
 S40111  
 capsid protein - human calicivirus (strain Bristol isolate B493)







```
Matches 140; Conservative 58; Mismatches 170; Indels 194; Gaps 28;
QY 4 ASKATPSADGATGAGQOLVPEVNTADPIPDPVAGSSTALATAG-----QVNLDPWII 57
DB 1779 AGTATTASVGGTTDMDPGVVAITSVVTAEE---NSSASIAATAGIGGPPQVDQOETWRT 1835
QY 58 NNFWQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMRVRVVLAGNA 117
DB 1836 NFYY---NDVFTWSVADAPGSLYTVQHSPPQNPFTAVLSQMYAGWAGMQFRFVAGIG 1892
QY 118 FTAGVVICCVPPGFQ-SRTLSIAQATLPPHHVADVRTLDPEVPLEDVNRVLYH-NDNT 175
DB 1893 VGGRLVAIVPPGIEIGLEVRQ---FPHVVIDARSLEPTIITPDLRPNMYHTGDP 1949
QY 176 QTMRLCLMLYPLRTGGASGDTSFVWAGRVLTCPGDFNFLLVLPPTVEOKTRPFTVP 235
DB 1950 GLVPTLVLSVYNNL-INPPGSGTSAIQVT--VETPSEDFFVMI----- 1991
QY 236 NIPLKYLNSRIPNPIEGMSLSPDQTONVQFQNGRCTTDGQPLGTPV---SVSOLCKFRG 293
DB 1992 -----RAPSSKTVDISF-----AGLLTTPVLTVGVGNDNRWG 2024
QY 294 RITSGORVLNLTDLGSPFMAFAAPAGFPDLGSCDWHIEMSKIPNSSTQNNPIVTNSV 353
DB 2025 QIVGLQ-----VPGGF-----STCNRRHNL-----NGSTY----- 2051
QY 354 KPNSQQFVPHLSSITLDENVSSGGDYIGTIQWTS-----PSDGGGANTN---F 399
DB 2052 -----WSSPRFGDIGHRRGSASYPGNNAITNVQLF 2080
QY 400 WKIPDYGSLAE-ASQAPAVYPPGFNEVIVFMASIP--GP-----NQS 441
DB 2081 W-YANAGSAIDNPISQVAPDGPDP-----MSFVPPNGPGIPAAAGWVGFAGIWSNS 2130
QY 442 GSPNLVPCLLPOEYITHFISEQAPIQGEAALLHYVDPTNRNLGFEKLYPGCYLTCVPN- 500
DB 2131 GAPNVTTV---QAYELGP-ATGAP-----GNLQPTNTS-----GSQTVAKSI 2169
QY 501 -SSSTGPGQPLDGVVFVFSWV 521
DB 2170 YAVVTGTAGNPA-GLFVWASGV 2190

RESULT 9
C48562
coat protein - San Miguel sea lion virus (serotype 4)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: C48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of San Miguel sea
eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: C48562
A:Molecule type: genomic RNA
A:Residues: 1-703 <NEI>
A:CROSS-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888
A>Note: sequence extracted from NCBI backbone (NCBIP:113567)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.9%; Score 285; DB 1; Length 703;
Best Local Similarity 24.1%; Pred. No. 3.2e-12;
Matches 116; Conservative 77; Mismatches 194; Indels 94; Gaps 23;
QY 12 ADGATGAGOLVPE---VNTADPIPDPVAGSSTALATAGVNLDPMI-----INN 59
DB 153 SDG-PGSAEIVTEEGTVVQQQAPAPA-PTALATLATAGTK-SVQEWNTTFFSYHTSIN- 208
QY 60 FVQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMRVRVVLAGNAPT 119
```

```
DB 209 -----WSTVESQKILYSQALNPSINPYLDHIAKLYSTWGGIDVRFRTVSGGVF 258
QY 120 AKGVICCVPPGFQ-SRTLSIAQATLPPHHVADVRTLDPEVPLEDVNRVLYHNDTOPT 178
DB 259 GGLAALLVPPGVPIESVSMQ---YHVLFDARQTEPVIPTDIRKTLPHSMDTDT 315
QY 179 MRLCLMLYPLRTGGASGDTSFVWAGRVLTCPGDFNFLLVLP--TVEQKTRP-TVP 235
DB 316 TKL---VINPYENGVENKTTCSITVETR-----PSADFTFALLKPPGSLIKHGISPSDLIP 368
QY 236 NIPLKYLNSRIPNPIEGMSLSPDQTONVQFQNGRCTTDGQPLG-TTPVVSUOLCKFRG 294
DB 369 RNSAHMGN-RWMSTISGFSVQPRVFS-----NRHFDSDSTTTGWTSTPYVPIETKIOCK 423
QY 295 IITSGORVLNLTDLGSPFMAFAAPAGFPDLGSCDWHIEMSKIPN--SSTQNNPIVTNS 352
DB 424 VGSNNKFWHVDTD---KALVPGIPDQWPD-----TTIPDETAKTNGNFSYGES 469
QY 353 VKPNSQQFVPHLSSITLDENVSSGGDYI-GTIQWTSPPSDSGGANTNFWKIPDYGSSLA 411
DB 470 YRAGSTTIKPN-----ENSTHFKGTICGLSTVEIPEN-----DEQIKTE 511
QY 412 ASQAPAVYPPGFNEVIVFMASIPGPNQSGSPNLVPCLLPOEYITHFISEQAPIQGEA 471
DB 512 AEKKSQWY-----VVTADFKDTIVKPOHK-----ISPQKLVYFDGPEKDLTMSAT 558
QY 472 L 472
DB 559 L 559

RESULT 10
A48562
coat protein - San Miguel sea lion virus (serotype 1)
N:Alternate names: capsid protein
C:Species: San Miguel sea lion virus
C>Date: 17-Feb-1994 #sequence_revision 17-Feb-1994 #text_change 23-Jul-1999
C:Accession: A48562
R:Neill, J.D.
Virus Res. 24, 211-222, 1992
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea
eins.
A:Reference number: A48562; MUID:92410750; PMID:1529644
A:Accession: A48562
A:Molecule type: genomic RNA
A:Residues: 1-702 <NEI>
A:CROSS-references: GB:M87481; NID:g334882; PIDN:AAAL6217.1; PID:g334884
A>Note: sequence extracted from NCBI backbone (NCBIP:113564, NCBIP:113565)
C:Superfamily: feline calicivirus coat protein
C:Keywords: capsid protein; coat protein; glycoprotein
F:208,481,493,545/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.7%; Score 281; DB 1; Length 702;
Best Local Similarity 21.8%; Pred. No. 6.1e-12;
Matches 116; Conservative 85; Mismatches 219; Indels 112; Gaps 21;
QY 2 MMASKDATP-----SADGATGAGOLVPEVNTA---DPIDIPVAGSSTALATAGVNL 52
DB 137 LQSTVDTKPDVFRRESGDPGADIVTEEGTVVQQQPV---PAQSALTTLAASGTGTV 193
QY 53 DPWIIINPVQAPQGEFTI-----SPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNG 102
DB 194 D-----CEWTTTFSSVHTAVNWSSTEAQSKILFSRALSPELNPYLRRHISLYST 241
QY 103 WGNMRVRVLAGNAFTAGKVIICVPPGFSQRTLSIAQATL--PPHVIAVTRTLDPEV 160
DB 242 WSGGIDVRFRTVSGGVFGGLAALIVPPGIE-----PVESPTMLQYPHYLFARQTEPVI 297
QY 161 PLEDVRNVLVYHNDTOPTMRLLCLMLYPLRTGCGASGDTSFVWAGRVLTTCGPDNPLFL 220
DB 298 TIPDIRKTLXHSMDTDTTTRLVIMVYNELINPYEOSEPKS-SCSITVETPSSDFTFSL 356
```



A:Accession: C45538  
A:Molecule type: Genomic RNA; protein  
A:Residues: 1-671 <CAR2>  
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
A:Experimental source: strain F9  
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)  
R:Gulver, M.; Littler, E.; Caul, E.O.; Fox, A.J.  
J. Gen. Virol. 73, 2429-2433, 1992  
A:Title: The cloning, sequencing and expression of a major antigenic region from the feline calicivirus (FCV) strain F9  
A:Reference number: PQ0407; MUID:93019069; PMID:1402818  
A:Accession: PQ0407  
A:Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>  
A:Cross-references: PIDN:AAB23553.1; PID:g257083  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.1%; Score 264.5; DB 1; Length 671;  
Best Local Similarity 33.3%; Pred. No. 8.3e-11;  
Matches 73; Conservative 32; Mismatches 89; Indels 25; Gaps 7;

QY 13 DCATGAGQLVPEVNT-ADPDPVAGSSTALATAGQVNLDPWIINNFWQAPQGEFTI- 70  
DB 127 DGSITA-----PQGTWVGVIAPSAQMSAADMATGKSDSEW-----EAFSFSFH 173

QY 71 -----SPNPTGDLVFLDLQGLPHLPFLSHLSQMYNGVNNRVVWLAGNAFTAGKVII 125  
DB 174 TSVNMTSETQKILFKQSLGPLLPYLEHLAKLVANSGSIEVRFSGSGVFGKLA 233

QY 126 CCVPPGQFS-RTLSIAQATLPHVIADVRLDPEVPLEVNRVLYHNNDTQPTMRLLCM 184  
DB 234 IVVPPGVPVQSTMLQ---YHVLFDARQVEPVIFCLPDLRSTLYHLMSDTRDTTSLVIM 290

QY 185 LYTPLRTGASGTSFVVGAGVLTCPGDFNLFVLP 223  
DB 291 VYNDLINPYANDANSSGCIV-TVETKPGDFKFKHLLKPP 328

RESULT 14  
VCWVFF  
coat protein - feline calicivirus (strain CFI/68 FIV)  
N:Alternate names: capsid protein  
C:Species: feline calicivirus  
C:Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text\_change 20-Aug-1999  
C:Accession: A40507; B40507; T09246  
R:Neill, J.D.; Reardon, I.M.; Heinrikson, R.L.  
J. Virol. 65, 5440-5447, 1991  
A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calicivirus (FCV) strain CFI/68  
A:Reference number: A40507; MUID:91374597; PMID:1716692  
A:Accession: A40507  
A:Molecule type: genomic RNA  
A:Residues: 1-668 <NE1>  
A:Cross-references: GB:M32819; NID:g323874; PIDN:AAA42925.1; PID:g323875  
A:Accession: B40507  
A:Molecule type: protein  
A:Residues: 373-379; 403-419; 481-489; 560-566 <NE2>  
R:Neill, J.D.  
submitted to the EMBL Data Library, April 1998  
A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68.  
A:Reference number: Z16626  
A:Accession: T09246  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: Genomic RNA  
A:Residues: 1-668 <NE3>  
A:Cross-references: EMBL:U13992; NID:g3056875; PIDN:AAC13993.1; PID:g537256  
A:Experimental source: strain CFI/68 FIV  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.1%; Score 263.5; DB 1; Length 668;

Best Local Similarity 25.9%; Pred. No. 9.7e-11;  
Matches 105; Conservative 59; Mismatches 162; Indels 79; Gaps 18;

QY 23 PEVNT-ADPDPVAGSSTALATAGQVNLDPWIINNFWQAPQGEFTI-----SPNPT 75  
DB 133 PEQGTWVGVIAPSAQMSAADMATGKSDSEW-----EAFSFSHTSVNMTSET 183

QY 76 PGDVLFDLQGLPHLPFLSHLSQMYNGVNNRVVWLAGNAFTAGKVII CCVPPGQFS- 134  
DB 184 QKILFKQSLGPLLPYLEHLAKLVANSGSIVDRFSISGSGVFGKLAII VVPPGIDPV 243

QY 135 RTLSIAQATLPHVIADVRLDPEVPLEVNRVLYHNNDTQPTMRLLCMLYPLRTGGA 194  
DB 244 QSTSMQ---YHVLFDARQVEPVIFCLPDLRSTLYHLMSDTRDTTSLVIMVNDLINPYA 300

QY 195 SGGTDSFVVGAGVLTCPGDFNLFVLPPTVEQTRPTVPNIPILKLSNS-----RI 247  
DB 301 NDSNSSGCIV-TVETKPGDFKFKHLLKPP-----GSMLTGHSIPSLIPKSSSLWGNRP 354

QY 248 PNPIEGMSLSP---DQTNVQFN---QRCITDGOPLGTTVPVSVSQLCKFRGRITSGORV 301  
DB 355 WSDITDFVIRPFVFOANRHFDFNQBTAGWSTPRFRPI-TITISVKESAKLGIGVATDVI 413

QY 302 LNLTELDGSPMAFAPAGFPDLGSCDHEIEMSKIPNSSTQNNPI-----VTNSVKNS 357  
DB 414 PGI-----PDGMPD-----TTIPGELV---PVGDIYAITNGTNDI 445

QY 358 QQFVPHLSSITLDENVSSGGDIY-GTIO--WTSPSPDSGGANTNF 399  
DB 446 TTAQYDAAIRNNTRFRGMYICGSLQRAW----GDKKISNTAF 486

RESULT 15  
JQ2356  
capsid protein - feline calicivirus (strain KCD)  
C:Species: feline calicivirus  
C:Date: 28-Aug-1985 #sequence\_revision 07-Oct-1994 #text\_change 20-Sep-1999  
R:Seal, B.S.; Ridpath, J.F.; Mengeling, W.L.  
J. Gen. Virol. 74, 2519-2524, 1993  
A:Title: Analysis of feline calicivirus capsid protein genes: identification of variable regions  
A:Reference number: JQ2354; MUID:94065683; PMID:7504075  
A:Accession: JQ2356  
A:Molecule type: mRNA  
A:Residues: 1-668 <SEA>  
A:Cross-references: GB:L09719; NID:g305107; PIDN:AAA16487.1; PID:g305108  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein

Query Match 8.8%; Score 253; DB 2; Length 668;  
Best Local Similarity 32.7%; Pred. No. 5.3e-10;  
Matches 70; Conservative 32; Mismatches 90; Indels 22; Gaps 7;

QY 19 GOLV-PEVNT-ADPDPVAGSSTALATAGQVNLDPWIINNFWQAPQGEFTI----- 70  
DB 128 GSIIVTPEQGTWVGVIAPSAQMSAADMATGKSDSEW-----EAFSFSHTSVNM 178

QY 71 SPNPTGDLVFLDLQGLPHLPFLSHLSQMYNGVNNRVVWLAGNAFTAGKVII CCVPP 130  
DB 179 STSETQKILFKQSLGPLLPYLEHLAKLVANSGSIEVRFSGSGVFGKLAII VVPP 238

QY 131 GFQS-RTLSIAQATLPHVIADVRLDPEVPLEVNRVLYHNNDTQPTMRLLCMLYPL 189  
DB 239 GVDPVQSTSMQ---YHVLFDARQVEPVIFCLPDLRSTLYHLMSDTRDTTSLVIMVNDL 295

QY 190 RTGASGTSDFVVGAGVLTCPGDFNLFVLP 223  
DB 296 INPYANDANSSGCIV-TVETKPGSDFRFHLKPP 328

Search completed: June 1, 2004, 13:55:22  
Job time : 12.4966 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (C) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.57181 Seconds  
(without alignments)  
4310.252 Million cell updates/sec

Title: US-09-926-799-4

Perfect score: 2891

Sequence: 1 MMASKDTPSADGATGAGQ.....YOLKPVGTAGPARGLGVR 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2234.5	77.3	546	1 COAT_SOUV3	Q04542 southampton
2	1156	40.0	539	1 COAT_LORDV	P54635 lordsdale v
3	306	10.6	2344	1 POLN_RHDV	P27410 rabbit hemo
4	285	9.9	703	1 COAT_SMSV4	P36285 san miguel
5	281	9.7	702	1 COAT_SMSV1	P36284 san miguel
6	278	9.6	2208	1 POLN_WANCV	Q69014 manchester
7	274	9.5	668	1 COAT_FCVP4	P27405 feline cali
8	264.5	9.1	671	1 COAT_FCVP9	P27406 feline cali
9	263.5	9.1	668	1 COAT_FCVC6	P27404 feline cali
10	148	5.1	2194	1 POLG_HE701	P32537 h genome po
11	135	4.7	2206	1 POLG_POL3L	P03302 poliovirus
12	134	4.6	2175	1 POLG_BOVEV	P12915 b genome po
13	132.5	4.6	2206	1 POLG_POL32	P06209 poliovirus
14	123	4.3	915	1 A180_RAT	Q05140 rattus norv
15	122	4.2	2205	1 POLG_POL2W	P23069 p genome po
16	122	4.2	2207	1 POLG_POL2L	P06210 p genome po
17	121.5	4.2	2194	1 POLG_EC30B	Q9W78 e genome po
18	119.5	4.1	1233	1 IRS1_MOUSE	P35569 mus musculu
19	119	4.1	2153	1 POLG_HRV16	Q82122 h genome po
20	117.5	4.1	1235	1 IRS1_RAT	P35570 rattus norv
21	115.5	4.0	855	1 POLG_HRV3	Q82081 human rhino
22	115.5	4.0	2179	1 POLG_HRV14	P03303 h genome po
23	115	4.0	2196	1 POLG_EC05N	Q9V1J1 e genome po
24	114	3.9	862	1 POLG_EC16H	Q66790 echovirus 1
25	113.5	3.9	760	1 SIX4_HUMAN	Q9U1U6 homo sapien
26	113.5	3.9	2333	1 POLG_FMDV1	P03306 f genome po
27	113	3.9	1075	1 NFK3_HUMAN	Q12968 homo sapien
28	113	3.9	2150	1 POLG_HRV2	P04936 h genome po
29	113	3.9	2206	1 POLG_CXA21	P22055 c genome po
30	112.5	3.9	2193	1 POLG_HE71B	Q66478 h genome po
31	112	3.9	907	1 A180_HUMAN	Q60641 homo sapien
32	112	3.9	2208	1 POLH_POL1M	P03300 p genome po
33	111.5	3.9	896	1 CYRB_MOUSE	P26955 mus musculu

34 111.5 3.9 897 1 CYRB\_HUMAN P32927 homo sapien  
35 111 3.8 832 1 POLG\_HRV1A P23008 human rhino  
36 111 3.8 2185 1 POLG\_SVDVH P16604 s genome po  
37 111 3.8 2185 1 POLG\_SVDVU P13900 s genome po  
38 110.5 3.8 769 1 TME8\_MOUSE Q9esn3 mus musculu  
39 110 3.8 520 1 LAC1\_AGABI Q12541 agaricus bi  
40 110 3.8 521 1 RUN2\_HUMAN Q13950 h runt-rela  
41 110 3.8 607 1 RUN2\_MOUSE Q08775 m runt-rela  
42 109 3.8 2209 1 POLG\_POL1S P03301 p genome po  
43 108.5 3.8 1024 1 SZ6L\_HUMAN Q9byh1 homo sapien  
44 108 3.7 2193 1 POLG\_CX16T Q9qf31 c genome po  
45 107 3.7 1324 1 SALL1\_HUMAN Q9nsc2 homo sapien

## ALIGNMENTS

### RESULT 1

ID COAT\_SOUV3 STANDARD; PRT; 546 AA.  
AC Q04542;  
DT 01-OCT-1994 (Rel. 30, Created)  
DT 01-OCT-1994 (Rel. 30, Last sequence update)  
DT 01-OCT-1994 (Rel. 30, Last annotation update)  
DE Coat protein (Capsid protein).  
OS Southampton virus (serotype 3).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=37129;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=93142023; PubMed=8380940;  
RA Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;  
RT "Sequence and genome organization of a human small round-structured (Norwalk-like) virus."  
RL Science 259:516-519(1993).  
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
EMBL; L07418; AAA92984.1; -  
DR PIR; B37491; B37491.  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
KW Coat protein; Glycoprotein.  
FT CARBOHYD 73 73 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 303 303 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 340 340 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 441 441 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;

Query Match 77.3%; Score 2234.5; DB 1; Length 546;

Best Local Similarity 74.9%; Pred. No. 2.8e-150;

Matches 408; Conservative 58; Mismatches 78; Indels 1; Gaps 1;

OY 1 MMASKDTPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLDPVNNF 60  
1 MMASKDTPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLDPVNNF 60  
Db 1 MMASKDTPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLDPVNNF 60  
61 VOAPQGETTISPNNTPGDVLFDLQLGPHLNPFLSHLSQMYNGWGMVRVVLNAGNFTA 120  
61 VQSPQGETTISPNNTPGDVLFDLQLGPHLNPFLSHLSQMYNGWGMVRVVLNAGNFTA 120  
QY 121 GKVIICVPPGFGQSRTLSIAQATLPPHVIADVRLDPVEVPLEDVRNVLVYHNDQPTMR 180  
121 GKIIIVCCVPPGFTSSSLTIAQATLPPHVIADVRLDPVEVPLEDVRNVLVYHNDQPTMR 180

		Matches	257; Conservative	78; Mismatches	172; Indels	62; Gaps	17;
QY	1	MMASKADTPSADGATGACQQLVPEVNTADPIPDVAGSSTALATAGQWNLDPMIINNF	60				
Db	1	MMASKADTPSADGATGACQQLVPEVNTADPIPDVAGSSTALATAGQWNLDPMIINNF	56				
QY	61	VQAPQCEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA	120				
Db	57	VQAPQCEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA	116				
QY	121	GVIIICVPPGFSRSLTAQATLPHVITADVRLDVEPLEDVRNVLVH--NNDTQPTM	179				
Db	117	GVIIICVPPGFSRSLTAQATLPHVITADVRLDVEPLEDVRNVLVH--NNDTQPTM	176				
QY	180	RLLCMLYTLPLRTGASGGTDSFVWAGRVLTCPGDFNLFVLPVPTVEQKTRPFTVPNIP	239				
Db	177	KLIALMLYTLPLRTGASGGTDSFVWAGRVLTCPGDFNLFVLPVPTVEQKTRPFTVPNIP	234				
QY	240	KYLSNSRIPNPIEGMSLSPDQTONVQFNGRCTIDQPLGTTTPSVSOLCKRGRIT--	296				
Db	235	EEMSNSRFPPIPLEKLYTGFSFAFVWPQNGRCTIDGVLGTTQLSAVINICFRGDVTHIA	294				
QY	297	-SGORVNLTELDGSP-MAFAAPAPAGPD--LGSCDWHIEMSKIENSSTQNNPIVNSV	353				
Db	295	GSHTYTNLWLASQNSWSDPTPEEIPALGTPDFVKGKIQGLLTQTTRADGSTRAKA--TV	351				
QY	354	KPSQOQVPHLSITLDENVSSGGDYIGTIQWTSPPSD--SGGANTNF	399				
Db	352	STGSHVTFPKL-----GSVQFTDNTNDFQAGQNTKFTPVGVIOGDH	395				
QY	400	-----WKIPDYSSSLAEASQAPAVYPPGFNEVIVYFMAISIPGNSQSGPNL-VPCLLP	452				
Db	396	QNEPQMSLNPYSGRITGHNVLAPAVPTFGQLLFFRSTMPG--CSGYPNMLDCLLP	453				
QY	453	QYIITHRISEAPIQGEAALLHYVDPDTNRNLAGEFKLYPGGYLTCVNSSTGQQ--LP	510				
Db	454	QEWLHFYQEAAPQASDVALLRFNVDPTGRVLFCKLHKSQYITV---AHTGPDVLVP	509				
QY	511	LDGVFVFASWSRFFYQLKPVGTAGPARGR	539				
Db	510	PNGYFRFDSWVWQFVTLAPMGN-GTGRRR	537				
RESULT 3							
ID	POLN RHDV	STANDARD;	PRT;	2344	AA.		
AC	P27410;						
DT	01-AUG-1992 (Rel. 23, Created)						
DT	01-AUG-1992 (Rel. 23, Last sequence update)						
DT	10-OCT-2003 (Rel. 42, Last annotation update)						
DE	Non-structural polyprotein [Contains: RNA-directed RNA polymerase						
DE	(EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like						
DE	protein); Coat protein].						
OS	Rabbit hemorrhagic disease virus (RHDV).						
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;						
OC	Lagovirus.						
OX	NCBI_TaxID=11976;						
RN	SEQUENCE FROM N.A.						
RP	MEDLINE=91361557; PubMed=1840711;						
RX	MEDLINE=91361557; PubMed=1840711;						
RA	Meyers G., Wirblich C., Thiel H.-J.;						
RT	"Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide						
RT	sequencing of a calicivirus genome.";						
RL	Virology 184:664-676(1991).						
CC	- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE						
CC	FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.						
CC	- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +						
CC	(RNA) (N).						
CC	- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration						
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -						
CC	the European Bioinformatics Institute. There are no restrictions on its						



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M67473; AAA47285.1; -  
 DR PIR; A41039; BRWHEH  
 DR PDB; 1KHV; 16-JAN-02.  
 DR MEROPS; C24.001; -  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR004005; Calici coat.  
 DR InterPro; IPR004004; Calici\_pol\_hel.  
 DR InterPro; IPR000317; Peptidase\_C24.  
 DR InterPro; IPR000605; RNA helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_Psvir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 DR Pfam; PF03510; Peptidase\_C24; 1.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA helicase; 1.  
 DR PRINTS; PR00916; 2CENDOPTASE.  
 DR PRINTS; PR00918; CALICIVIRUS.  
 DR SMART; SM00382; AAA; 1.  
 KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;  
 KW Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.  
 FT CHAIN 7508 7707 HELICASE (P2C-LIKE).  
 FT CHAIN 1109 1251 PROTEASE P3C.  
 FT CHAIN 71497 71625 RNA-DIRECTED RNA POLYMERASE.  
 FT CHAIN ? 2344 COAT PROTEIN.  
 FT NP\_BIND 522 529 ATP (POTENTIAL).  
 FT ACT\_SITE 1212 1212 PROTEASE (POTENTIAL).  
 FT ACT\_SITE 1227 1227 PROTEASE (POTENTIAL).  
 SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 10.6%; Score 306; DB 1; Length 2344;

Best Local Similarity 24.2%; Pred. No. 2.4e-13;

Matches 147; Conservative 66; Mismatches 184; Indels 210; Gaps 30;

QY 4 ASKDATPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAG-----QVNLDPWII 57  
 DB 1779 AGTATTASVPGTTGMDPGVATTSVITAE---NSSASITAGIGGPPQDQOETWRT 1835  
 QY 58 NNFVQAPQCEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSOMYNGWGNRVVRVLGNA 117  
 DB 1836 NFYY---NDVFTWSVADAFSGSILYTVQHSPPQNNPFTAVLSQMYAGWAGQGFVAGSG 1892  
 QY 118 FTAGKVIIICVPPGFQ-SRTLIAQATLPPHVIADVRLDPVEVPLEDVRLVLYH-NNDT 175  
 DB 1893 VFGRLVRAVIPPGEIGPGLVRQ---PPHVVIDARSLEPVTITMPDLRPNMYHTGDP 1949  
 QY 176 QPTMELLCMLYPLRTGGASGTDSPVAVAGRLVTCGPDFNPLFLVPPVVEQKTRPFTVP 235  
 DB 1950 GLVPTLVLSVNNL-INPPGSGTSAIQVT--VETRPSEDFEFVMIRAPSSK----- 1997  
 QY 236 NIPLKYLNSRIPNPIEGMSLPDQTONVQNGRCTIDG-OPLG--TTPV--SVSQLCK 290  
 DB 1998 -----TVDSISPAGLLITPVLVTGVGNDR 2021  
 QY 291 FGRITSGORVLNLTLDGSPFMAFAAPAGFPDLGSCDMHIEMSKIPNSSTNNPIVT 350  
 DB 2022 WNGQIVGLQ-----PVPGGF---STCNRHWNL-----NGSTYG----- 2051  
 QY 351 NSVKNSQOFVPHLSITLDENVSSGGDYIGTIQWTSP-----PSDGGANTV- 398  
 DB 2052 -----WSSPRFGIDHRRGSASYSGSNATV 2077  
 QY 399 --FMKIPDYGSIAE-ASQAPAVPPGPNFVVFMAISIP--GP----- 438  
 DB 2078 LQFW-YANAGSAIDNPISQVADPGFPD-----MSFVPPNGGIPAGWVGAIWN 2127

QY 439 NOSGPNLVPCLLPQYIYTHF-----ISEQAPIQGEAALLHYVDPDTRN-LGSP 487  
 DB 2128 SNSGAPNTTV---QAYELGFATGAPGNLQPTTNTSGAQTVAKSIYAVVVTGAQNPAGLP 2184  
 QY 488 KLYPGGYLTCVPSNST-----GPQQLPL--DGVFVFASWVSFPYQLKPVG 531  
 DB 2185 VMASG--IISTPNASAITYPQPDRIVTPGTPAAAPVKGKTPIMFASVVRRTGDVN--A 2240  
 QY 532 TAGPARG 538  
 DB 2241 TAGSANG 2247  
 RESULT 4  
 COAT\_SMSV4 STANDARD; PRT; 703 AA.  
 AC P36285;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS San Miguel sea lion virus (serotype 4) (SMSV 4).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus  
 OC NCBI\_TaxID=36407;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92410750; PubMed=1529644;  
 RA Neill J.D.;  
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of  
 RT San Miguel sea lion viruses: identification of conserved and non-  
 RT conserved amino acid sequences among calicivirus capsid proteins.";  
 RL Virus Res. 24:211-222 (1992).  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M87482; AAA16220.1; -

PIR; C48562; C48562.

InterPro; IPR004005; Calici\_coat.

InterPro; IPR008975; Viral\_cap\_coat.

Pfam; PF00915; Calici\_coat; 1.

Coat protein; Glycoprotein.

FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 703 AA; 77721 MW; CSDAD8223B261073 CRC64;

Query Match 9.9%; Score 285; DB 1; Length 703;

Best Local Similarity 24.1%; Pred. No. 1.4e-12;

Matches 116; Conservative 77; Mismatches 194; Indels 94; Gaps 23;

QY 12 ADGATGAGQLVPE-----VNTADPIIDPVAGSSTALATAGQVNLDPWI-----INN 59

DB 153 SDG-PGSAEIVTEEGTVVQQQPAPA-PTALATLATASTGK-SVEQWMTFVSHTSIN- 208

QY 60 FVQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSOMYNGWGNRVVRVLGNAFT 119

DB 209 -----WSTVESQKILYSQALNPSINPYLDHIAKLYSTVSGGIDVRETVSGGVF 258

QY 120 AGKVIICVPPGFQ-SRTLIAQATLPPHVIADVRLDPVEVPLEDVRLVLYHNDTQPT 178

DB 259 GGLKLAALLVPPGVEPIESVMLQ----YHVLFDARQTEPVIPTIDIRKTLFHSMDTDT 315

QY 179 MRLCMLYPLRTGGASGTDSPVAVAGRLVTCGPDFNPLFLVPP--TVEQKTRPF-TVP 235

316 TKL---VINPYGVENKTCISITVETR-----PSADFTFALLKPPGSLIKHGSIPSDLIIP 368  
 236 NPLKYLNSRIPNPIEGMSLSPDQTONVQFONGRECTIDGQPLG-TTPVSVSQLCKRGR 294  
 369 RNSAHWMGN-RWWSITSGFSVQPRVFQS-----NRHFDFTSTTGTGMSVPIPIEIKIGK 423  
 295 IFSQORVLNLTDLGSPFMAFAAPAGPFDLGSCDWHIEMSKIPN--SSTQNNPIVITNS 352  
 424 VGSNNKWFVIDTD-----KALVPGIDGWD-----TTIDETKATNGNFSYGES 469  
 353 VKPNSQOQFPHLSIITDENSSGGDYI-GTIQWTSPPSDSGGANTNFWKIPDIDYSSLA 411  
 470 YRAGSTTIKEN-----ENSTFKGTICGTLSTVEIPEN-----DEQIKTE 511  
 412 ASQAPAVTPPGNEVIVFMASIPGPNQSGSNLVPCLLPOEYITHFISEQAPQGEAA 471  
 512 AEKKSQTM-----VVTADFQKOTIVKPOHK-----ISPQKLVVYFDGPEKDLTMSAT 558  
 472 L 472  
 559 L 559

RESULT 5

COAT SMSV1 STANDARD; PRT; 702 AA.  
 AC P36284; 1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS San Miguel sea lion virus (serotype 1) (SMSV 1).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesiviruses.  
 OX NCBI\_TaxID=36406;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=92410750; PubMed=1529644;  
 RA Neill J.D.;  
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of  
 RT San Miguel sea lion virus: identification of conserved and non-  
 RT conserved amino acid sequences among calicivirus capsid proteins.";  
 RL Virus Res. 24:211-222(1992).  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M87481; AAA16217.1; --  
 DR PIR; A48562; A48562.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR Pfam; PF008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DEED37 CRC64;

Query Match 9.7%; Score 281; DB 1; Length 702;  
 Best Local Similarity 21.8%; Pred. No. 2.8e-12;  
 Matches 116; Conservative 85; Mismatches 219; Indels 112; Gaps 21;  
 2 MMASKDATP-----SADGATGAGQLVPEVNTA---DPIPIPVAGSSTALATAGQVNL 52  
 137 LQSTVDTKDFVFRASDGGADIVTEBQGTVVQOQPV---PAQSALTTLAAASTGKTV 193

53 DPWIIINNFVQARQGBETI-----SPNNTPGDVLFDLQGLPHLNFPLSHLSQMNG 102  
 194 D-----CEWTTTFSTHTAVNWSSTEAQKILFSRALSPELNPYLRIHSSLYST 241  
 103 WUGNMRVRVVLAGNAFTAGKVIICCPGFGQRTLSIAQATL--FPHVIADVRTLDPEV 160  
 242 WSGGIDVRFVTVSGSGVFGKLAALIVPPGIE-----PVESPTMLQYPHVLFDARQTEPVIF 297  
 161 PLEBVRNVLVHNNDTPTWRLICMLTYPLRTGASGSDTSFVVAGRVLTCPGDFNFLL 220  
 298 TIPDIRKTLVHSMDDTDTTTLVIMVYNELINPYEAGEPKS-SCSITVETRPSSDFTFSL 356  
 221 VPTVEQKTRPFTVPNIP--LKYLNSRIPNPIEGMSLSPDQTONVQFONGRECTIDGQPL 278  
 357 KPGSLLKHGSI PDSILPNSRHHMGNRWWSITDGFVQPRVFQS-----NRHFDFTSTTT 412  
 279 GTT-----PVSVSQCKFRGRITSGQRVLNLTDLGSPFMAFAAPAGPFDL----- 326  
 413 GWSTPYVPIEVT-----LEKLDGQGYFKVDTDEKS-----LVPGLPDGDWEDTTIFAMT 463  
 327 ---GSCDWHIEMSKIPNSSTQNNPIVITNSVKPNSQOQFVPHLSSITLDENVSS---GGDYI 380  
 464 ASNGNYDYTTVAEYRITNNGT-----HFKGYIMGNLTTKVKGSDNL 504  
 381 GTIQWTSPPSDSGGANTNFWKIPDYSSLAESAQAPAVYPPGFNEVIVFMASIPGPNQ 440  
 505 GETQOTS-----RTLFAVGNY-----KDQNTINP-THKITSNSLVVYDANNV---S 547  
 441 SGSPNLVPCLLPOEYITHFISEQAPQGEAALLHYVD--PDTNRNLGEPKLY 490  
 548 AATARTTTHSHTMHGLGVLDSPGSDTKVVRITATPEATNGNGNFFV 599

RESULT 6

POLN MANCV STANDARD; PRT; 2208 AA.  
 ID POLN MANCV  
 AC Q69014;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Genome polyprotein [Contains: RNA-directed RNA polymerase  
 DE (EC 2.7.7.48); thiol protease 3C (EC 3.4.22.-); Helicase (2C like  
 DE protein); Coat protein].  
 OS Manchester virus (Human enteric calicivirus).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Sapovirus.  
 OX NCBI\_TaxID=82659;  
 [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=95390791; PubMed=7661689;  
 RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;  
 RT "Human enteric caliciviruses have a unique genome structure and are  
 RT distinct from the Norwalk-like viruses.";  
 RL Arch. Virol. 140:1345-1356(1995).  
 CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).  
 CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR  
 CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY  
 CC SIMILARITY).  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA}(N).  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----



ID COAT\_FCVC6 STANDARD; PRT; 671 AA.  
 AC DT27406;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Coats protein (Capsid protein).  
 DE 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Coats protein (Capsid protein).  
 GN CYP76.  
 OS Feline calicivirus (strain F9) (FCV).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OC NCBI\_TaxID=11981;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=92410623; PubMed=1529544;  
 RX Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,  
 RA Turner F.C.;  
 RA "The complete nucleotide sequence of a feline calicivirus";  
 RT Virology 190:443-448(1992).  
 RL [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=92339673; PubMed=1633955;  
 RX Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;  
 RA "Cloning and sequence determination of the feline calicivirus strain  
 F9";  
 RT Biochem. Soc. Trans. 20:26S-26S(1992).  
 RL [1]  
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC [2]  
 DR EMBL; M86379; AAA79327.1; -;  
 DR EMBL; Z11536; CAA77636.1; -;  
 DR PIR; B43382; VCMWF9.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 DR Coats protein; Glycoprotein.  
 KW CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT SEQUENCE 671 AA; 73441 MW; 33BEE86D8370D5ES CRC64;  
 SQ  
 Query Match 9.1%; Score 264.5; DB 1; Length 671;  
 Best Local Similarity 33.3%; Pred. No. 3.8e-11;  
 Matches 73; Conservative 32; Mismatches 89; Indels 25; Gaps 7;  
 QY 13 DGATGAGQLVPEVNT-ADPDPVAGSSTALATAGQVNLDPMLINNFVQAPQGEFTI- 70  
 Db 127 DGSITA-----PQGTWVGVIAPSAQNSAADMATGKSVDSW-----EAFSFSFH 173  
 QY 71 -----SPNPTPDVLFDLQGLPHLPFLSHLSQMYNGVGNRVVVLGNAGNAGTAKGVII 125  
 Db 174 TSVNWSSTSETQKILFKQSLGSLPFLPYLTHLAKLVANSGSVDFRFSISGSGVFGKGLAA 233  
 QY 126 CCVPRGFQOS-RTLSTAQATLFFPHVIADVRLTPVEVPLEDVRNVLVHNNDTQPTMRLLCM 184  
 Db 234 IVPVPGVDVQSTSMQLQ---YHVLFDARQVEVIFCLPDLRLSTLYHLMSDITDTTSLVIM 290  
 QY 185 LYTPLRTGAGSGTDSFVVGAVRLVTCGPDENFLFLVPP 223  
 Db 291 VYNDLINPANDANSNGCIV-TVETKPGDPDFKHLKPP 328  
 RESULT 9  
 COAT\_FCVC6 STANDARD; PRT; 668 AA.  
 ID -COAT\_FCVC6

AC DT27404;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Coats protein (Capsid protein).  
 DE 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Coats protein (Capsid protein).  
 GN CYP76.  
 OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OC NCBI\_TaxID=11979;  
 OX [1]  
 RN SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RP MEDLINE=91374597; PubMed=1716692;  
 RX Neill J.D., Reardon I.M., Heinrikson R.L.;  
 RA "Nucleotide sequence and expression of the capsid protein gene of  
 RT feline calicivirus";  
 RL J. Virol. 65:5440-5447(1991).  
 RN [2]  
 RN SEQUENCE FROM N.A.  
 RP Neill J.D.;  
 RA Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 CC -1- PTM: The N-terminus is blocked.  
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC [2]  
 DR EMBL; M32819; AAA42925.1; -;  
 DR EMBL; U13992; AAC13993.1; -;  
 DR PIR; A40507; VCMWF9.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 DR Coats protein; Glycoprotein.  
 KW CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT SEQUENCE 668 AA; 73550 MW; 9E52312108D4D441 CRC64;  
 SQ  
 Query Match 9.1%; Score 263.5; DB 1; Length 668;  
 Best Local Similarity 25.9%; Pred. No. 4.4e-11;  
 Matches 105; Conservative 59; Mismatches 162; Indels 79; Gaps 18;  
 QY 23 PEVNT-ADPDPVAGSSTALATAGQVNLDPMLINNFVQAPQGEFTI-----SPNPT 75  
 Db 133 PQGTWVGVIAPSAQNSAADMATGKSVDSW-----EAFSFSFHTSVNWSSTSET 183  
 QY 76 PGDVLFDLQGLPHLPFLSHLSQMYNGVGNRVVVLGNAGNAGTAKGVIIICVPPGFOS- 134  
 Db 184 QKILFKQSLGSLPFLPYLTHLAKLVANSGSVDFRFSISGSGVFGKGLAAIVVPPGIDPV 243  
 QY 135 RTLSTAQATLFFPHVIADVRLTPVEVPLEDVRNVLVHNNDTQPTMRLLCMLYTLPLRTGGA 194  
 Db 244 QSTSMQLQ---YHVLFDARQVEVIFSPDLRLSTLYHLMSDITDTTSLVIMVNDLINPYA 300  
 QY 195 SGTDSFVVGAVRLVTCGPDENFLFLVPPTVBQKTRPFTVPNIPKLYLSNS-----RI 247  
 Db 301 NDSNNGCIV-TVETKPGDPDFKHLKPP-----GSMTHGSIPLSLIPKSSSLWGNRF 354  
 QY 248 PNPFGMSLSP---DOTQNVQFQN---GRCTIDGQPLGTPFVSVQOLCKFRGRIISGQVR 301  
 Db 355 WSDITDFVIRPFVQANRHFDFNQETAGWSTFRFPI-TITISVKESAKGLGIVATYIV 413  
 QY 302 LNITELDGGSPFAFAAPAGPDLGSCDWHIEMSKIPNSSTONNPI-----VTNSVKPNS 357  
 Db 414 PGI-----PDGMPD-----TTIPGLV---PVGYAITNGTNDI 445

QY 358 QQVPHLSITLDENVSSGDYI-GTIO--WTSPPSDSCGANTNF 399  
 DB 446 TTAQYDAATEIRNTNFRGMVIGSLQRAW-----GDKKISNTAF 486

## RESULT 10

POLG\_HE701  
 ID POLG\_HE701 STANDARD; PRT: 2194 AA.  
 AC P32537;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
 DE Human enterovirus 70 (strain J670/71) (EV 70).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.  
 OC NCBI\_TaxID=31915;  
 RX MEDLINE=91037960; PubMed=2172447;  
 RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D., Minor P.D., Almond J.W.;  
 RT "The complete nucleotide sequence of enterovirus type 70: relationships with other members of the picornaviridae.";  
 RL J. Gen. Virol. 71:2291-2299 (1990)  
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).  
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DB EMBL; D00820; BAA18891.1; --  
 DB PIR; A36253; GNNYE7.  
 DB HSP; P03300; IPOV.  
 DB MEROPS; C03.UBA; --  
 DB MEROPS; C03.UBA; --  
 DR InterPro; IPR003593; AAA ATPase.  
 DR InterPro; IPR004004; Calici\_pol\_hel.  
 DR InterPro; IPR009003; Cys\_ser\_trypsin.  
 DR InterPro; IPR000199; Pept\_3C\_picorn.  
 DR InterPro; IPR000081; Peptidase\_C3.  
 DR InterPro; IPR001338; Pico\_P1A.  
 DR InterPro; IPR002527; Pico\_P2B.  
 DR InterPro; IPR001676; Rhv.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_PSVir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00548; Cys-protease-3C; 1.  
 DR Pfam; PF02226; Pico\_P1A; 1.  
 DR Pfam; PF00947; Pico\_P2A; 1.  
 DR Pfam; PF01552; Pico\_P2B; 1.

DR Pfam; PF00073; rhv; 3.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PD00918; CALICVIRUSNS.  
 DR ProDom; PD001125; Cys\_protease\_3C; 1.  
 DR ProDom; PD001306; Pico\_P2A; 1.  
 DR ProDom; PD001274; Pico\_P2B; 1.  
 DR SMART; SM00382; AAA; 1.  
 KW Polyprotein; Coat protein; Core protein; Transferase;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 KW Lipoprotein.  
 FT CHAIN 2 69 COAT PROTEIN VP4.  
 FT CHAIN 70 319 COAT PROTEIN VP2.  
 FT CHAIN 320 561 COAT PROTEIN VP3.  
 FT CHAIN 552 871 COAT PROTEIN VP1.  
 FT CHAIN 872 1014 CORE PROTEIN P2A.  
 FT CHAIN 1015 1113 CORE PROTEIN P2B.  
 FT CHAIN 1114 1443 CORE PROTEIN P2C.  
 FT CHAIN 1444 1532 CORE PROTEIN P3A.  
 FT CHAIN 1533 1737 GENOME-LINKED PROTEIN VP3.  
 FT CHAIN 1555 1737 PICORNAIN 3C.  
 FT CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.  
 FT LIPID 2 N-myristoyl glycine (in host) (By similarity).  
 FT ACT\_SITE 1701 1701 PROTEASE (POTENTIAL).  
 FT ACT\_SITE 1715 1715 PROTEASE (POTENTIAL).  
 SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;  
 Query Match 5.1%; Score 148; DB 1; Length 2194;  
 Best Local Similarity 22.1%; Pred. No. 0.031;  
 Matches 91; Conservative 53; Mismatches 163; Indels 104; Gaps 19;  
 QY 57 INNVQAPQGE---FTISPNNTFGDVLF-----DLQL-GPHLPFLSHLSQMYNGWVGNMR 108  
 DB 374 INNVNDASGVERLRVQISAQSDMDQLFNIPLDIQLEGLPLNTLLGNISRYTHWSGLE 433  
 QY 109 VRVVLAGNAFTAGKVIICVPPGFSQRTLSIAQATLPHVIADVRLDPEVPLEDVNV 168  
 DB 434 MTFMFCGSMFTTKLIICVTPPGSSPT-DRMQAMLATHVVMDFGLQSSITIIIPISGS 492  
 QY 169 LVH--NNDTQPMR---LLCMLYPLRTGGASGSDTSFVAGRV-----LTCPCP 213  
 DB 493 HYRMNTAKAINANVGYTCFMQNL--VAPVGAADQCYIVGMVAKKDFNRLMRDSP 550  
 QY 214 DNFNLFVLPVPP-----TVEQ--KTRPFTVPNIPLKYLNSRIPNPEGMS 255  
 DB 551 DIGQSAILPEQAATQIGEVKTVANTVESEIKALGVIPSLNAVETGATSTEEPAI- 609  
 QY 256 LSPDQTNVQFONG--RCTIDGQPIGTTPSVSVQCKPGRITSGQRV-----LNLT 305  
 DB 610 ----QTRTVNNHGTAECLVENF-LGRSALVCMRSFEYKNHSTSTSSIQKNFFIWTNTR 664  
 QY 306 EL-----DGSFFMAF-----AAPAGAPDPLGSCDWHIENSKIENSTQNNPIVTSV 353  
 DB 665 ELVQIRKRMELFTYLRFDTEITIVTLRLFSS-----NVSGLENLTQAWYVPTGAR 719  
 QY 354 KPNSQVFPHLSSITLDENVSSGGDYIGTIOVTPSPDSGGANTNFWKIPD 404  
 DB 720 KPSSQD-----SFEWQACNFS-----VFFKIND 743  
 RESULT 11  
 POLG\_POL3L  
 ID POLG\_POL3L STANDARD; PRT: 2206 AA.  
 AC P03302; Q84783; Q84784; Q84785; Q84786; Q84787; Q84788; Q84789;  
 AC Q84790; Q98592; Q98593; Q98594;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C, P3A; Genome-linked protein VP3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].

Wed Jun 2 09:13:29 2004

OS Poliovirus type 3 (strains P3/Leon/37 and P3/Leon 12A[1]B).  
OC Viruses; serNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Enterovirus.  
OX NCBI\_TaxID=12088;  
RN (1)  
RN SEQUENCE FROM N.A.  
RP STRAIN=P3/Leon/37;  
RX MEDLINE=84170338; PubMed=6324200;  
RA Stanway G., Hughes P.J., Mountford R.C., Reeve P., Minor P.D.,  
RA Schild G.C., Almond J.W.;  
RT "Comparison of the complete nucleotide sequences of the genomes of  
RT the neurovirulent poliovirus P3/Leon/37 and its attenuated Sabin  
RT vaccine derivative P3/Leon 12A[1]B."  
RL Proc. Natl. Acad. Sci. U.S.A. 81:1539-1543 (1984).  
RN (2)  
RN SEQUENCE FROM N.A.  
RP STRAIN=P3/Leon 12A[1]B;  
RX MEDLINE=83292339; PubMed=6310508;  
RA Stanway G., Cann A.J., Hauptmann R., Hughes P.J., Clarke L.D.,  
RA Mountford R.C., Minor P.D., Schild G.C., Almond J.W.;  
RT "The nucleotide sequence of poliovirus type 3 Leon 12 alb: comparison  
RT with poliovirus type 1."  
RL Nucleic Acids Res. 11:5629-5643 (1983).  
RN (3)  
RN X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS) OF 1-878.  
RP MEDLINE=95120467; PubMed=7820548;  
RX Grant R.A., Hiramath C.N., Filman D.J., Syed R., Andries K.,  
RA Hogle J.M.;  
RT "Structures of poliovirus complexes with anti-viral drugs:  
RT implications for viral stability and drug design."  
RL Curr. Biol. 4:784-797 (1994).  
RN (4)  
RN X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 1-878.  
RP Hiramath C.N., Grant R.A., Filman D.J., Hogle J.M.;  
RA "Binding of the antiviral drug wint51711 to the Sabin strain of type-3  
RT poliovirus - structural comparison with drug-binding in rhinovirus-  
RT 14."  
RL Acta Crystallogr. D 51:473-489 (1995).  
CC -1- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
CC O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the  
CC poliovirus polypeptide. In other picornavirus reactions Glu may be  
CC substituted for Gln, and Ser or Thr for Gly.  
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA}(N).  
CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
CC each of which is composed of one copy each of proteins VP1, VP2,  
CC VP3, and VP4.  
CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC -1- MISCELLANEOUS: THE SEQUENCE OF STRAIN SABIN VACCINE P3/LEON/37 IS  
CC SHOWN.  
CC -1- MISCELLANEOUS: THE STRAIN SABIN VACCINE P3/LEON/37 IS THE  
CC PROGENITOR OF THE STRAIN SABIN VACCINE P3/LEON 12A[1]B.  
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; K01392; AAA46914.1; -;  
CC EMBL; X00925; CA25444.1; -;  
CC PDB; 1PTV; 03-JUN-95.  
CC PDB; 1PVC; 15-SEP-95.  
CC PDB; 1VBA; 11-JUL-96.  
CC PDB; 1VBB; 11-JUL-96.  
CC PDB; 1VBC; 11-JUL-96.  
CC PDB; 1VBE; 11-JUL-96.  
CC MEROPS; C03.001; -;  
CC MEROPS; C03.020; -;  
DR InterPro; IPR004004; Calici\_pol\_hel.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR000199; Peptidase\_C3.  
DR InterPro; IPR000081; Peptidase\_C3.  
DR InterPro; IPR003138; Pico\_P1A.  
DR InterPro; IPR002527; Pico\_P2B.  
DR InterPro; IPR001676; Rhv.  
DR InterPro; IPR000605; RNA\_helicase.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00548; Cys-protease\_3C; 1.  
DR Pfam; PF02226; Pico\_P1A; 1.  
DR Pfam; PF00947; Pico\_P2A; 1.  
DR Pfam; PF01552; Pico\_P2B; 1.  
DR Pfam; PF00073; rhv; 3.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
DR PRINTS; PR00918; CALICIVIRUSNS.  
DR PRODOM; PD001125; Cys\_protease\_3C; 1.  
DR PRODOM; PD001306; Pico\_P2A; 1.  
DR PRODOM; PD001274; Pico\_P2B; 1.  
DR Polyprotein; Coat protein; Core protein; Transferase;  
DR RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
KW 3D-structure; Lipoprotein.  
FT CHAIN 2 69  
FT CHAIN 70 340  
FT CHAIN 341 578  
FT CHAIN 579 878  
FT CHAIN 879 1027  
FT CHAIN 1028 1124  
FT CHAIN 1125 1453  
FT CHAIN 1454 1540  
FT CHAIN 1541 1562  
FT CHAIN 1563 1745  
FT CHAIN 1746 2206  
FT LIPID 2 2  
FT ACT SITE 1709 1709  
FT ACT\_SITE 1723 1723  
FT VARIANT 431 431  
FT VARIANT 864 864  
FT VARIANT 908 908  
FT STRAND 4 8  
FT STRAND 25 29  
FT STRAND 36 38  
FT TURN 50 50  
FT TURN 51 54  
FT STRAND 57 57  
FT TURN 63 64  
FT TURN 80 81  
FT STRAND 83 87  
FT TURN 88 89  
FT STRAND 90 94  
FT STRAND 101 102  
FT STRAND 103 105  
FT TURN 113 115  
FT STRAND 123 123  
FT STRAND 126 128  
FT TURN 129 129  
FT STRAND 133 134  
FT STRAND 138 141  
FT TURN 142 143  
FT STRAND 147 151  
FT TURN 152 152  
FT STRAND 153 155  
FT TURN 156 157  
FT STRAND 159 167  
FT STRAND 168 180  
FT TURN 185 185  
FT STRAND 187 197  
FT TURN 198 198  
FT STRAND 203 203  
COAT PROTEIN VP4.  
COAT PROTEIN VP2.  
COAT PROTEIN VP3.  
COAT PROTEIN VP1.  
CORE PROTEIN P2A.  
CORE PROTEIN P2B.  
CORE PROTEIN P2C.  
CORE PROTEIN P3A.  
GENOME-LINKED PROTEIN VPG.  
PICORNAIN 3C.  
RNA-DIRECTED RNA POLYMERASE P3D.  
N-myristoyl glycine (in host).  
PROTEASE (POTENTIAL).  
PROTEASE (POTENTIAL).  
S -> F (IN P3/LEON 12A[1]B).  
K -> R (IN P3/LEON 12A[1]B).  
T -> A (IN P3/LEON 12A[1]B).

```

FT TURN 208 209
FT HELIX 213 216
FT STRAND 219 221
FT STRAND 223 224
FT STRAND 226 226
FT TURN 235 236
FT STRAND 242 242
FT STRAND 245 245
FT HELIX 246 248
FT TURN 249 252
FT HELIX 255 260
FT STRAND 263 267
FT TURN 268 270
FT STRAND 273 278
FT STRAND 287 287
FT TURN 289 291
FT STRAND 292 292
FT STRAND 295 307
FT TURN 308 309
FT STRAND 310 310
FT STRAND 314 331
FT STRAND 337 337
FT TURN 348 351
FT TURN 355 356
FT STRAND 363 363
FT TURN 366 367
FT STRAND 379 380
FT STRAND 382 382
FT TURN 383 383
FT HELIX 384 387
FT TURN 388 388
FT STRAND 391 392
FT STRAND 397 397
FT TURN 399 403
FT HELIX 405 408
FT STRAND 410 413
FT TURN 414 415
FT TURN 418 419
FT STRAND 423 426
FT TURN 429 431
FT TURN 433 437
FT HELIX 439 444
FT TURN 445 446
FT STRAND 447 451

Query Match
Best Local Similarity 4.7%; Score 135; DB 1; Length 2206;
Matches 61; Conservative 30; Mismatches 110; Indels 50; Gaps 10;

QY 22 VPEVNTADPIPDVAGSSTALATAGQVNLIDPWVQAPQG-----EFTISPNNT 76
Db 365 IPEFDVTPDIPDGEVKNMELAE-----IDTWIPLNLESTKNTMDMYRVTLSADL 418
QY 77 GNVFLDLQGLPHNLFSLH-----LSQMTNGVGNMVRVVLVAGNAFTAGKVIICVPPG 131
Db 419 SQPILCLSPASDRLSHTMLGELVNLVYTHWAGSLKFTFLFCGSMATGKILVAYAPG 478
QY 132 FQSRTLSIAQTLFPHVIADVTLDPEVPLEDVRNVLVHNNDTPTMR---LLCWLXY-- 186
Db 479 AQPT-SRKEALMGLTHVIMDLGLQSSCTVWVWVSNVYTRQT-TQDSFTGGYISMFYQT 536
QY 187 ---TPLRTGASGGTDSFVYVAGRVLTCPGDFNFLLFPVPTVEQKTRPTVFNILPKLV 243
Db 537 RIVVPLST-----PKSMGLGFSVAC--NDFSRLRLRDTT-----HIS 572
QY 244 NSRINPIEGM 254
Db 573 QSALPQGIEDL 583

RESULT 12
POLG_BOVEV STANDARD; PRT; 2175 AA.
ID POLG_BOVEV

```

```

AC P12915;
DT 01-OCT-1989 (Rel. 12, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
DE linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].
OS Bovine enterovirus (strain VG-5-27) (BEV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=12065;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88117392; PubMed=2828511;
RA Earle J.A.P., Skuce R.A., Fleming C.S., Hoey E.M., Martin S.J.;
RT "The complete nucleotide sequence of a bovine enterovirus.";
RL J. Gen. Virol. 69:253-263(1988).
RN [2]
RP X-RAY CRYSTALLOGRAPHY (3.0 ANGSTROMS) OF 2-840.
RX MEDLINE=95292108; PubMed=7773791;
RA Smyth M., Tate J., Hoey E.M., Lyons C., Martin S.J., Stuart D.;
RT "Implications for viral uncoating from the structure of bovine
RT enterovirus.";
RL Nat. Struct. Biol. 2:224-231(1995).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the
CC poliovirus polypeptide. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; D00214; BAA24003.1; ALT_SEQ.
CC PIR; A29824; GNYYBE.
CC DR PDB; 1BEV; 16-SEP-98.
CC DR MEROPS; C03.UBA; -.
CC DR MEROPS; C03.UPB; -.
CC DR InterPro; IPR004004; Calici_pol_hel.
CC DR InterPro; IPR009003; Cys_Ser_trypsin.
CC DR InterPro; IPR001199; Pept_3C_picorn.
CC DR InterPro; IPR000081; Peptidase_C3.
CC DR InterPro; IPR003138; Pico_P1A.
CC DR InterPro; IPR002527; Pico_P2B.
CC DR InterPro; IPR001876; Rhv.
CC DR InterPro; IPR006005; RNA_helicase.
CC DR InterPro; IPR007095; RNA_pol_PS.
CC DR InterPro; IPR001205; RNA_pol_P3D.
CC DR InterPro; IPR007094; RNA_pol_P5vir.
CC DR InterPro; IPR008975; Viral_cap_coat.
CC DR Pfam; PF00548; Cys-protease_3C; 1.
CC DR Pfam; PF02226; Pico_P1A; 1.
CC DR Pfam; PF00947; Pico_P2A; 1.
CC DR Pfam; PF01552; Pico_P2B; 1.
CC DR Pfam; PF00073; rhv; 3.
CC DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC DR Pfam; PF00910; RNA_helicase; 1.
CC DR PRINTS; PD00918; CALICVIRUSN.
CC DR ProDom; PD001125; Cys_protease_3C; 1.

```

DR ProDom; PD001306; Pico\_P2A; 1.  
DR ProDom; PD001274; Pico\_P2B; 1.  
KW Polypeptide; Coat protein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
KW 3D-structure; Lipoprotein.  
FT CHAIN 2 69  
FT CHAIN 70 317  
FT CHAIN 318 559  
FT CHAIN 560 840  
FT CHAIN 841 990  
FT CHAIN 991 1089  
FT CHAIN 1090 1419  
FT CHAIN 1420 1508  
FT CHAIN 1509 1531  
FT CHAIN 1532 1714  
FT CHAIN 1715 2175  
FT CHAIN 2176 2175  
FT CHAIN 2175 1678  
FT ACT SITE 1678 1678  
FT ACT SITE 1692 1692  
FT STRAND 26 26  
FT STRAND 36 38  
FT TURN 50 50  
FT HELIX 51 54  
FT STRAND 57 57  
FT STRAND 83 87  
FT TURN 88 89  
FT STRAND 90 94  
FT STRAND 101 102  
FT HELIX 103 105  
FT TURN 113 115  
FT STRAND 123 123  
FT HELIX 126 128  
FT TURN 129 129  
FT STRAND 133 134  
FT STRAND 138 139  
FT TURN 142 143  
FT STRAND 147 151  
FT TURN 152 152  
FT HELIX 153 155  
FT TURN 156 157  
FT HELIX 159 167  
FT STRAND 168 180  
FT TURN 185 186  
FT STRAND 187 197  
FT TURN 198 198  
FT HELIX 209 212  
FT STRAND 215 217  
FT STRAND 219 220  
FT HELIX 224 226  
FT TURN 227 230  
FT STRAND 234 238  
FT STRAND 241 245  
FT TURN 246 248  
FT STRAND 251 256  
FT STRAND 265 265  
FT TURN 267 269  
FT STRAND 270 270  
FT STRAND 273 285  
FT STRAND 294 309  
FT TURN 325 328  
FT TURN 332 333  
FT STRAND 340 340  
FT TURN 343 344  
FT STRAND 356 357  
FT STRAND 359 360  
FT TURN 360 360  
FT HELIX 361 364  
FT TURN 365 365  
FT STRAND 368 369  
FT STRAND 375 375  
FT TURN 376 377  
FT TURN 380 381  
FT HELIX 382 384

FT STRAND 389  
FT TURN 394  
FT STRAND 397  
FT TURN 408  
FT HELIX 410  
FT STRAND 413  
FT TURN 421  
FT STRAND 423  
FT STRAND 430  
FT TURN 440  
FT STRAND 443  
FT STRAND 445  
FT HELIX 461  
FT TURN 466  
FT STRAND 468  
FT STRAND 479  
FT TURN 493  
FT STRAND 499  
FT HELIX 503  
FT STRAND 510  
FT STRAND 520  
FT TURN 523  
FT STRAND 528  
FT TURN 539  
FT STRAND 542  
FT TURN 550  
FT STRAND 574  
FT STRAND 578  
FT STRAND 582  
FT TURN 595  
FT STRAND 597  
FT HELIX 610  
FT STRAND 619  
FT STRAND 624  
FT HELIX 626  
FT STRAND 629  
FT HELIX 630  
FT STRAND 638  
FT STRAND 644  
FT TURN 647  
FT STRAND 651  
FT HELIX 662  
FT TURN 669  
FT STRAND 671  
FT STRAND 687  
FT TURN 689  
FT STRAND 693  
FT STRAND 705  
FT TURN 707  
FT HELIX 718  
FT STRAND 727  
FT TURN 732  
FT TURN 733

Query Match 4.6%; Score 134; DB 1; Length 2175;  
Best Local Similarity 18.4%; Pred. No. 0.3;  
Matches 88; Conservative 57; Mismatches 161; Indels 172; Gaps 17;  
31 IPIDVAGSSTALAT-----AGQVNLDPWLNINNVQAPQGEFTISPNNTPG- 77  
319 LPTKPGSGYQFWMTTDEDCSPCILPFOPTLEIFIPGKVNLLLETAQVESILEANNREGV 378  
78 -----DVLFDLQL-----GPHLNPFLSHLSOMYNGWGNMVRVULA 114  
379 EGVRYVIVSVQDAL-DAQIYALRLGSGPLSSLLGLTAKHVTQWSGSEITCMFT 437  
115 GNAFTAGKVIICCPFGFQSRTLSIAQATLFPHVIADVTLPDVEVPL----- 162  
438 GTFTWTKGVLLAYTPFG-GDMPNREAMLGTHVWDFGLQSSITLVIPIWISASHPRGVS 496  
163 -EDVRNVLHNNDTQPTMRLLCMLYTPLTGGASGDTDFWVAGRVLTCPGPDF----- 215  
497 NDDVLNYYQY-----NAGHTIYQTNWVIP---PGFPNTAGI 531



QY 216 -----NFLFLVPTVQKTRPFTVPNIPKYLNSRIPNPIEGM-----SUS 257  
 Db 532 IMMIAAQPNSFRIDKREDMTQTAILQNDPGKMLKDA-IDKQVAGALVAGTTTSTHVA 590  
 QY 258 PPOQTQNVQFN-----GCTIDGSLGTTTPVS 284  
 Db 591 TOSTPALQAAETGATSTARDESMIETRTIVPTHGIIHETSVESFFGSSLVGMPLIATGTS 650  
 QY 285 VS-----QLCKFRGRI-----TSQRVNLNLTDLGSPFMAFAAPA 319  
 Db 651 IINWRIDREFVQLRAKMSFTYMRDVEFTIIATSTGQ---NVTTEQHTTYQWYVPP 707  
 QY 320 PAGFP-DLQSCDWHIEMSKIPNSSTONNPIVNSKPNQSQFVPHLSITLBNVSSG 376  
 Db 708 GAFVPSNQDSFQW-----QSGCNPSVFADTDGPPAQFSVPFMSANAYSTVYDG 756

RESULT 13  
 POLG\_POL32  
 ID POLG\_POL32 STANDARD; PRT: 2206 AA.  
 AC P06209;  
 DT 01-JAN-1988 (Rel. 06, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Coat proteins VP1 TO VP4; Core proteins P2A TO P2C, P3A; Genome-linked protein VP3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
 OS Poliovirus type 3 (strain 23127).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 OX NCBI\_TaxID=12087;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=87010550; PubMed=3020156;  
 RA Hughes P.J., Evans D.M.A., Minor P.D., Schild G.C., Almond J.W., Stanway G.;  
 RT "The nucleotide sequence of a type 3 poliovirus isolated during a recent outbreak of poliomyelitis in Finland.";  
 RL J. Gen. Virol. 67:2093-2102(1986).  
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA}(N).  
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: X04468; CAA28155.1; -.  
 DR PIR: A27245; GNNV27.  
 DR HSRP: Q84730; LPVC.  
 DR MEROPS: C03.001; -.  
 DR MEROPS: C03.020; -.  
 DR InterPro: IPR003593; AAA ATPase.  
 DR InterPro: IPR004004; Calici\_pol\_hel.  
 DR InterPro: IPR009003; Cys\_ser\_tpsin.  
 DR InterPro: IPR000199; Pept\_3C\_picorn.  
 DR InterPro: IPR000081; Peptidase\_C3.  
 DR InterPro: IPR003138; Pico\_P1A.

DR InterPro: IPR002527; Pico\_P2B.  
 DR InterPro: IPR001676; Rhv.  
 DR InterPro: IPR000605; RNA\_helicase.  
 DR InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro: IPR001205; RNA\_pol\_P3D.  
 DR InterPro: IPR007094; RNA\_pol\_P5vir.  
 DR InterPro: IPR008975; Viral\_cap\_coat.  
 DR Pfam: PF00548; Cys-protease-3C; 1.  
 DR Pfam: PF02226; Pico\_P1A; 1.  
 DR Pfam: PF00947; Pico\_P2A; 1.  
 DR Pfam: PF01552; Pico\_P2B; 1.  
 DR Pfam: PF00073; rhv; 3.  
 DR Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam: PF00910; RNA\_helicase; 1.  
 DR PRINTS; PR00918; CALICVIRUSNS.  
 DR ProDom; PD001125; Cys\_protease\_3C; 1.  
 DR ProDom; PD001306; Pico\_P2A; 1.  
 DR ProDom; PD001274; Pico\_P2B; 1.  
 DR SMART; SM00382; AAA; 1.  
 KW Polyprotein; Coat protein; Core protein; Transferase;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 KW Lipoprotein.  
 FT CHAIN 2 69 COAT PROTEIN VP4.  
 FT CHAIN 70 340 COAT PROTEIN VP2.  
 FT CHAIN 341 578 COAT PROTEIN VP3.  
 FT CHAIN 579 878 COAT PROTEIN VP1.  
 FT CHAIN 879 1027 CORE PROTEIN P2A.  
 FT CHAIN 1028 1124 CORE PROTEIN P2B.  
 FT CHAIN 1125 1453 CORE PROTEIN P2C.  
 FT CHAIN 1454 1540 CORE PROTEIN P3A.  
 FT CHAIN 1541 1562 GENOME-LINKED PROTEIN VP3.  
 FT CHAIN 1563 1745 PICORNAIN 3C.  
 FT CHAIN 1746 2206 RNA-DIRECTED RNA POLYMERASE P3D.  
 FT LIPID 2 2 N-mristoyl glycine (in host) (By similarity)  
 FT ACT\_SITE 1709 1709 PROTEASE (POTENTIAL).  
 FT ACT\_SITE 1723 1723 PROTEASE (POTENTIAL).  
 SQ SEQUENCE 2206 AA; 245731 MW; F226AD85403C7BA CRC64;  
 Query Match 4.6%; Score 132.5; DB 1; Length 2206;  
 Best Local Similarity 24.4%; Pred No. 0.39;  
 Matches 63; Conservative 37; Mismatches 121; Indels 37; Gaps 13;  
 QY 22 VPEVNTADPIPDPVAGSSTALATAGQVNLDPWIIINNFVQAPGSEFTI---SPNNTPGD 78  
 Db 365 IPEPDVTPPIDI---PGEVKNVMELEIDTMIPLNLENTKRTMDMYRVRLSDSANTLSCP 421  
 QY 79 VLFDLQLGPHLNPFLSH-----LSQYNGVGNMVRVVLGNAGTAKVIICCVPPGQ 133  
 Db 422 ILC-LSLSPAADPRLSHTMLGEVLNYYTHWAGSLKFTFLFCGSMMATGKLLVAYAPPGAQ 480  
 QY 134 SRTLSIAQATLPPHVIADVRLTDPVEVPLEDVRNVLYHNNDTQPTMR---LLCMLY--- 186  
 Db 481 PPT-SRKAMGLTHVINDGLQSSCTMVVPWISNYYTQRT-TQDSFTGGYISMFYQTRI 538  
 QY 187 -TPLRTGASGGTDSFVVAGRVLTCPGDFNPLFLVPPT-VEQKTRPFTVPNIPKLYSN 244  
 Db 539 VVPLSTPRAMD-----MLGFVSAC--NDFSVRLLRDTTHISQAAMPQGVDDLITEVAQN 590  
 QY 245 S---RIPNPIEGMSLSPD 259  
 Db 591 ALALS LKPK---QSNLPD 605

RESULT 14  
 A180\_RAT  
 ID A180\_RAT STANDARD; PRT: 915 AA.  
 AC Q05140;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE Clathrin coat assembly protein AP180 (Clathrin coat associated protein AP180) (91 kDa synaptosomal-associated protein).  
 DE

GN SNAP91.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORMS LONG AND SHORT).  
 RC TISSUE=Brain;  
 RX MEDLINE=931178442; PubMed=8440257;  
 RA Morris S.A., Schroeder S., Flessmann U., Weber K., Ungewickell E.;  
 RT "Clathrin assembly protein AP180: primary structure, domain organization and identification of a clathrin binding site.";  
 RL EMBO J. 12:667-675(1993).  
 CC -!- FUNCTION: Adaptins are components of the adaptor complexes which link clathrin to receptors in coated vesicles. Clathrin-associated protein complexes are believed to interact with the cytoplasmic tails of membrane proteins, leading to their selection and concentration. Binding of AP180 to clathrin triskelia induces their assembly into 60-70 nm coats.  
 CC -!- SUBCELLULAR LOCATION: Component of the coat surrounding the cytoplasmic face of coated vesicles in the plasma membrane.  
 CC -!- ALTERNATIVE PRODUCTS:  
 CC -!- Event-Alternative splicing; Named isoforms=2;  
 CC Names=Long;  
 CC IsoId=Q05140-1; Sequence=Displayed;  
 CC Names=Short;  
 CC IsoId=Q05140-2; Sequence=VSP\_000173;  
 CC -!- DOMAIN: Possesses a three domain structure: the N-terminal 300 residues harbor a clathrin binding site, an acidic middle domain 450 residues, interrupted by an Ala-rich segment, and the C-terminal domain (166 residues).  
 CC -!- PTM: Phosphorylated (by similarity).  
 CC -!- SIMILARITY: Contains 1 epsilon N-terminal homology (ENTH) domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).  
 CC EMBL; X68877; CAA48748.1; -.  
 CC EMBL; X68878; CAA48749.1; -.  
 CC PIR; S36326; S36326.  
 CC DR; S36327; S36327.  
 CC InterPro; IPR001026; ENTH.  
 CC InterPro; IPR008943; PI\_bind\_N.  
 CC Pfam; PF01417; ENTH; 1.  
 CC SMART; SM00273; ENTH; 1.  
 CC PROSITE; PS50942; ENTH; 1.  
 CC Coated pits; Alternative splicing; Phosphorylation.  
 CC FT DOMAIN 14 145 ENTH.  
 CC FT DOMAIN 410 413 POLY-THR.  
 CC FT DOMAIN 535 539 POLY-ALA.  
 CC FT DOMAIN 547 550 POLY-ALA.  
 CC FT DOMAIN 678 683 POLY-SER.  
 CC FT DOMAIN 723 729 POLY-SER.  
 CC FT VARSPLIC 614 632 Missing (in isoform Short).  
 CC SEQUENCE 915 AA; 93518 MW; 32EC1B38EC5DF8C0 CRC64;  
 Query Match 4.3%; Score 123; DB 1; Length 915;  
 Best Local Similarity 19.6%; Pred. No. 0.56;  
 Matches 124; Conservative 64; Mismatches 226; Indels 218; Gaps 27;  
 4 ASKDATPSADGATGAGQLVPEVNTA-----DIPIDP-----VAGSS 40  
 363 AAAPVPPTGATGAWGDLGCELSAALSSVSPSEAPISDFAPFAPSPPTTTTTPASASASA 422  
 41 TALATAG--QVNLIDPWINNFVQAPQGEFTISPNNTPGDVLFDLQGLPHLPFTLSLSQ 98  
 423 TTAVTAATTEVDLFGDAFAASPGAEPAASEGATAPATPAPVAALDACSGNDFFAPS--- 479

QY 99 MYNGWGNRVVVLGNAGNFTAGKVIICCVPPGFSQRTLSIAQATLPHPHVIADVRTLDVP 158  
 DB 480 -----EGSAEAAPELDLFAMKP-----PETSAPVVTPTASTAPPV 514  
 QY 159 EYPLEEDVRNVLVYHNNDTQTMRLLCMLYPLRTGASGCTDSFVVAGRVLTG----RGPD 214  
 DB 515 PA-----TAPSPAPTAVATAATTTTAAAAAATTTATTSAAAAATAAAPALD 560  
 QY 215 -FNFLFLVPPTVEQKTRPTVPNIPKYLNSRINPIEGMSLSPDQTONVQFQNG----R 270  
 DB 561 IFGDLUFDSAPEAAASKPDVAPSIDL--FGTDASSPPRGASPVPESSLTADLLSGSGFH 618  
 QY 271 CTIDGQ--PL-----GTFPVSQOLCKFRGRTSGQRVNLUTELDGSPPMAFAA-- 317  
 DB 619 CAEDDRHVPLFTTAVDAPAAPSPASTASPAKAE---SSG-----VIDLFGDAFGSSASET 670  
 QY 318 -PAP-----AGFPDLGSCDWHIEMKIPNSGTONN-----HLSSTILDENVSSG 376  
 DB 671 QPAPQAVSSSSASADLLAGF--GS---FMAPSTTPTVTPAQNNLLQPNFEAFGTTPTS 725  
 QY 347 -----PIVTNSVKPNSQOFV-----HLSSTILDENVSSG 376  
 DB 726 SSSSFPDPSGDLMLPTWAPSGQAPVGMVPPSPAMSAKGLGSLDLSLASLVNGLIGIST 785  
 QY 377 GDYIGTIQWTSPPSD--SGGANTNFWK-----IPDYGSLAEASQALAPAVVPPGF 424  
 DB 786 TSKKGDGLQWNAKEKLTGGAN---WQPKVTPATWSAGVPPQGTVPPTSS-----VPPG- 835  
 QY 425 NEVIVYPMASIRPGNOSGSPNLVP-----CLLPQEVYTHIFISEQAIQGEAALLHYVDP 478  
 DB 836 -----AGAPSVGQGAGYGMPAGTGMTWMPQPPVMAFQPMMPRPFGAAAV----- 881  
 QY 479 DTNRNLGFKLYPGGYLTCPVNSSTGTPQQLP 510  
 DB 882 -----PGQLSPSPPTATQSPKPP 901  
 RESULT 15  
 POLG POL2W  
 ID POL2W STANDARD; PRT; 2205 AA.  
 AC P23069;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2 (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Picornain 2A (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].  
 DE Poliovirus type 2 (strain W-2).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 OC NCBI\_TaxID=12085;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=90155230; PubMed=2154539;  
 RA Pevear D.C., Oh C.K., Cunningham L.L., Calenoff M., Jubeit B.;  
 RT "Localization of genomic regions specific for the attenuated, mouse-adapted poliovirus type 2 strain W-2.";  
 RL J. Gen. Virol. 71:43-52(1990).  
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).  
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 31.6725 Seconds

(without alignments)  
5419.275 Million cell updates/sec

Title: US-09-926-799-4

Perfect score: 2891

Sequence: 1 MMASKDTPSADGATGAGQ.....YQLKPVGTAGPARGRLGVRR 544

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL.25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2891	100.0	544	12 Q9DU46	Q9du46 chiba virus
2	2883	99.7	544	12 Q9QTE7	Q9qte7 chiba virus
3	2868	99.2	544	12 Q91I85	Q91i85 human calic
4	2858	98.9	544	12 Q918C5	Q918c5 norwalk-like
5	2853	98.7	544	12 Q91EJ2	Q91ej2 human calic
6	2852	98.7	544	12 Q918D1	Q918d1 norwalk-like
7	2825	97.7	544	12 Q91V41	Q91v41 human calic
8	2812	97.3	531	12 Q8V7J5	Q8v7j5 norwalk-like
9	2792	96.6	544	12 Q91V36	Q91v36 human calic
10	2361.5	81.7	543	12 Q91V43	Q91v43 human calic
11	2347.5	81.2	543	12 Q8JXJ0	Q8jxj0 norwalk-like
12	2346.5	81.2	543	12 Q918B9	Q918b9 norwalk-like
13	2231.5	77.2	546	12 Q8JW44	Q8jw44 norwalk-like
14	2226.5	77.0	546	12 Q91V47	Q91v47 human calic
15	2218.5	76.7	544	12 Q8VA27	Q8va27 human calic
16	2096	72.5	530	12 Q83884	Q83884 norwalk vir

17	2093	72.4	530	12 Q9QT39	Q9qt39 norwalk-like
18	2059.5	71.2	540	12 Q9YS14	Q9ys14 norwalk vir
19	2054	71.0	530	12 Q83876	Q83876 norwalk vir
20	2051	70.9	517	12 Q9JH41	Q9jh41 norwalk-like
21	2036.5	70.4	540	12 Q8BCA5	Q8bca5 human calic
22	1998	69.1	539	12 Q8JV55	Q8jv55 norwalk-like
23	1989	68.8	545	12 Q8VA02	Q8va02 human calic
24	1988	68.8	541	12 Q91V42	Q91v42 human calic
25	1982	68.6	545	12 Q918C8	Q918c8 norwalk-like
26	1982	68.6	545	12 Q91W12	Q91w12 norwalk-like
27	1976.5	68.4	544	12 Q66418	Q66418 desert shie
28	1976	68.4	545	12 Q918C2	Q918c2 norwalk-like
29	1964	67.9	539	12 Q91V48	Q91v48 human calic
30	1960	67.8	545	12 Q91H12	Q91h12 human calic
31	1939	67.1	545	12 Q91V45	Q91v45 human calic
32	1895	65.5	543	12 Q8BCA3	Q8bca3 human calic
33	1328.5	46.0	522	12 Q8B558	Q8b558 bovine ente
34	1319.5	45.6	522	12 Q8B425	Q8b425 norwalk-like
35	1312.5	45.4	522	12 Q8B4Y9	Q8b4y9 norwalk-like
36	1309.5	45.3	522	12 Q8B422	Q8b422 norwalk-like
37	1306.5	45.2	520	12 Q7TBK7	Q7tbk7 bovine ente
38	1306.5	45.2	522	12 Q8V629	Q8v629 norwalk-like
39	1304.5	45.1	520	12 Q7TBL1	Q7tbl1 bovine ente
40	1301.5	45.0	522	12 Q8V628	Q8v628 norwalk-like
41	1268	43.9	519	12 Q9YQ22	Q9yq22 bovine cali
42	1192.5	41.2	556	12 Q9PY75	Q9py75 human calic
43	1192.5	41.2	556	12 Q91V76	Q91v76 norwalk-like
44	1184.5	41.0	556	12 Q91V79	Q91v79 norwalk-like
45	1180.5	40.8	542	12 Q8JMP1	Q8jmp1 norwalk-like

#### ALIGNMENTS

#### RESULT 1

Q9DU46 PRELIMINARY; PRT; 544 AA.  
AC Q9DU46;  
DT 01-MAR-2001 (Tremblrel. 16, Created)  
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Capsid protein.  
OS Chiba virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Calciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=99565;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Chiba 407/1987/JP;  
RA Someya Y., Takeda N., Miyamura T.;  
RL Submitted (MAY-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Chiba 407/1987/JP;  
RX MEDLINE=20569531; PubMed=11118371;  
RA Someya Y., Takeda N., Miyamura T.;  
RT "Complete nucleotide sequence of the Chiba virus genome and functional expression of the 3C-like protease in Escherichia coli.";  
RL Virology 278:490-500(2000).  
DR EMBL; AB042808; BAB18267.1;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 544 AA; 58351 MW; 1DB7A4E81FBB246F CRC64;

Query Match 100.0%; Score 2891; DB 12; Length 544;

Best Local Similarity 100.0%; Pred. No. 5.5e-225;

Matches 544; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MMASKDTPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLDPWIINF 60

DB 1 MMASKDTPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLDPWIINF 60

61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120  
121 GKVIICCVPPGQSRRTLSIAQATLPHVIADVRTLDPEVEPLEDRVNLVYHNNDTQPTMR 180  
121 GKVIICCVPPGQSRRTLSIAQATLPHVIADVRTLDPEVEPLEDRVNLVYHNNDTQPTMR 180  
181 LLCLMLYTPRLTGGASGTDTSFVAVAGRLVTCPPGDFNLFVLPPTVEQKTRPFTVPNIPK 240  
181 LLCLMLYTPRLTGGASGTDTSFVAVAGRLVTCPPGDFNLFVLPPTVEQKTRPFTVPNIPK 240  
241 YLSNSRIPNPIEGMSLSPDQTNQVQNGRCITDGOPLGTTTPVSVQSLCKFRGRTISQOR 300  
241 YLSNSRIPNPIEGMSLSPDQTNQVQNGRCITDGOPLGTTTPVSVQSLCKFRGRTISQOR 300  
301 VLNLTELDGSPFMAFAAPAPAGPDPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQOF 360  
301 VLNLTELDGSPFMAFAAPAPAGPDPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQOF 360  
361 VPHLSITLTDENVSSGGDYIGTIQWTSPPSDSGGANTFWKIPDYGSSLAELASQAPAVY 420  
361 VPHLSITLTDENVSSGGDYIGTIQWTSPPSDSGGANTFWKIPDYGSSLAELASQAPAVY 420  
421 PPGFNEVIVFMASIPGPNQSGSPNLVPCLLPOEYITHFISEQAPIQGEAALLHYVDPDT 480  
421 PPGFNEVIVFMASIPGPNQSGSPNLVPCLLPOEYITHFISEQAPIQGEAALLHYVDPDT 480  
481 NRNLGEFKLYPGGYLTCVPNSSSTGPOQLPDLGVFVFWFASVWSRFYQLKPVGTAGPARGRL 540  
481 NRNLGEFKLYPGGYLTCVPNSSSTGPOQLPDLGVFVFWFASVWSRFYQLKPVGTAGPARGRL 540  
541 GVR 544  
541 GVR 544

RESULT 3  
Q91185 PRELIMINARY; PRT; 544 AA.  
AC Q91185;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus HU/NLV/Koblenz/433/2000/DE.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=165252;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HU/NLV/Koblenz 433/2000/DE;  
RA Kuenkel U., Schreier E.;  
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany."  
RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF394960; RAK72048.1;  
DR InterPro: IPR004005; Calici coat.  
DR InterPro: IPR008975; Viral\_cap\_coat.  
DR Pfam: PF00915; Calici coat; 1.  
SQ SEQUENCE 544 AA; 58222 MW; 1B39A595B733A6A8 CRC64;

Query Match 99.2%; Score 2868; DB 12; Length 544;  
Best Local Similarity 99.1%; Pred. No. 3.9e-223;  
Matches 539; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 MMASKDTPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLDPWLNINNF 60  
DB 1 MMASKDTPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLDPWLNINNF 60  
QY 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120  
DB 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120  
QY 121 GKVIICCVPPGQSRRTLSIAQATLPHVIADVRTLDPEVEPLEDRVNLVYHNNDTQPTMR 180

61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120  
61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120  
121 GKVIICCVPPGQSRRTLSIAQATLPHVIADVRTLDPEVEPLEDRVNLVYHNNDTQPTMR 180  
121 GKVIICCVPPGQSRRTLSIAQATLPHVIADVRTLDPEVEPLEDRVNLVYHNNDTQPTMR 180  
181 LLCLMLYTPRLTGGASGTDTSFVAVAGRLVTCPPGDFNLFVLPPTVEQKTRPFTVPNIPK 240  
181 LLCLMLYTPRLTGGASGTDTSFVAVAGRLVTCPPGDFNLFVLPPTVEQKTRPFTVPNIPK 240  
241 YLSNSRIPNPIEGMSLSPDQTNQVQNGRCITDGOPLGTTTPVSVQSLCKFRGRTISQOR 300  
241 YLSNSRIPNPIEGMSLSPDQTNQVQNGRCITDGOPLGTTTPVSVQSLCKFRGRTISQOR 300  
301 VLNLTELDGSPFMAFAAPAPAGPDPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQOF 360  
301 VLNLTELDGSPFMAFAAPAPAGPDPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQOF 360  
361 VPHLSITLTDENVSSGGDYIGTIQWTSPPSDSGGANTFWKIPDYGSSLAELASQAPAVY 420  
361 VPHLSITLTDENVSSGGDYIGTIQWTSPPSDSGGANTFWKIPDYGSSLAELASQAPAVY 420  
421 PPGFNEVIVFMASIPGPNQSGSPNLVPCLLPOEYITHFISEQAPIQGEAALLHYVDPDT 480  
421 PPGFNEVIVFMASIPGPNQSGSPNLVPCLLPOEYITHFISEQAPIQGEAALLHYVDPDT 480  
481 NRNLGEFKLYPGGYLTCVPNSSSTGPOQLPDLGVFVFWFASVWSRFYQLKPVGTAGPARGRL 540  
481 NRNLGEFKLYPGGYLTCVPNSSSTGPOQLPDLGVFVFWFASVWSRFYQLKPVGTAGPARGRL 540  
541 GVR 544  
541 GVR 544

RESULT 2  
Q9QTE7 PRELIMINARY; PRT; 544 AA.  
AC Q9QTE7;  
DT 01-MAY-2000 (Tremblrel. 13, Created)  
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Capsid protein.  
OS Chiba virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=99565;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Chiba 407;  
RA Natori K., Suzuki K., Yamakawa Y., Tatsumi M., Sakae K., Kobayashi S., Shinozaki K., Ishiko H., Miyamura T., Takeda N.;  
RT "Expression and self-assembly of capsid proteins of the Chiba virus, a genetically distinct Norwalk-like virus."  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB022679; BAA82106.1;  
DR InterPro: IPR004005; Calici coat.  
DR InterPro: IPR008975; Viral\_cap\_coat.  
DR Pfam: PF00915; Calici coat; 1.  
SQ SEQUENCE 544 AA; 58323 MW; ABD1C1FC4F93D872 CRC64;

Query Match 99.7%; Score 2883; DB 12; Length 544;  
Best Local Similarity 99.6%; Pred. No. 2.4e-224;  
Matches 542; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MMASKDTPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLDPWLNINNF 60  
DB 1 MMASKDTPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLDPWLNINNF 60  
QY 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMVRVVLGNAFTA 120

```

Db      121  GKVILCCVPPGFSRSLTAQATLPHVIAVRTLDPEVEPLEDVNRVLYHNDTQPTNR 180
QY      181  LKMLYTLPLRTGASGGTDSFVAGRVLTCPGPDFNFLEFLVPTVEQKTRPPTVNPILK 240
Db      181  LKMLYTLPLRTGASGGTDSFVAGRVLTCPGPDFNFLEFLVPTVEQKTRPPTVNPILK 240
QY      241  YLSNSRIPNPIEGMSLSPDQTNVQFNGRCTIDGQPLGTTTPVSVSOLCKFRGRTISGQR 300
Db      241  YLSNSRIPNPIEGMSLSPDQTNVQFNGRCTIDGQPLGTTTPVSVSOLCKFRGRTISGQR 300
QY      301  VLNLTELDGSPFMAFAAPAGPDLGSCDWHIEMSKIPNSSTONNPIVTVNSVKNSQOF 360
Db      301  VLNLTELDGSPFMAFAAPAGPDLGSCDWHIEMSKIPNSSTONNPIVTVNSVKNSQOF 360
QY      361  VPHLSITLDDNVSSGGDIYGTIQTWSPPSDGGANTFWKIPDYGSSLAESAQLAPAY 420
Db      361  VPHLSITLDDNVSSGGDIYGTIQTWSPPSDGGANTFWKIPDYGSSLAESAQLAPAY 420
QY      421  PGFNEVIVYFMAIPGPNQSGSNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPT 480
Db      421  PGFNEVIVYFMAIPGPNQSGSNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPT 480
QY      481  NRNLGEFKLYPGGYLTCVPNSSTGQQLPLDGVFVFASWSRFFYQLKPVGTAGPARGRL 540
Db      481  NRNLGEFKLYPGGYLTCVPNSSTGQQLPLDGVFVFASWSRFFYQLKPVGTAGPARGRL 540
QY      541  GVR 544
Db      541  GVR 544

```

## RESULT 4

```

Q918C5  PRELIMINARY; PRT; 544 AA.
AC Q918C5;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Baltimore/277/1993/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171836;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Baltimore/277/1993/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577 (1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Baltimore/277/1993/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Baltimore/277/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348 (2000).
RN [4]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Baltimore/277/1993/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414404; AAL12965.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 544 AA; 58208 MW; D5170FA4B958A672 CRC64;

```

```

Query Match      98.9%; Score 2858; DB 12; Length 544;
Best Local Similarity 98.9%; Pred. No. 2.5e-222;
Matches 538; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MMASKDTPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGQVNLDPMIINNF 60
Db 1 MMASKDTPSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGVNLDPMIINNF 60
QY 61 VOAPQGEFTISNNTPGDLVFDLQGLPHLNPFLSHLSQMYNGWGNMRVRVVLGNAFTA 120
Db 61 VOAPQGEFTISNNTPGDLVFDLQGLPHLNPFLSHLSQMYNGWGNMRVRVVLGNAFTA 120
QY 121 GKVIICCVPPGFSRSLTAQATLPHVIAVRTLDPEVEPLEDVNRVLYHNDTQPTNR 180
Db 121 GKVIICCVPPGFSRSLTAQATLPHVIAVRTLDPEVEPLEDVNRVLYHNDTQPTNR 180
QY 181 LKMLYTLPLRTGASGGTDSFVAGRVLTCPGPDFNFLEFLVPTVEQKTRPPTVNPILK 240
Db 181 LKMLYTLPLRTGASGGTDSFVAGRVLTCPGPDFNFLEFLVPTVEQKTRPPTVNPILK 240
QY 241 YLSNSRIPNPIEGMSLSPDQTNVQFNGRCTIDGQPLGTTTPVSVSOLCKFRGRTISGQR 300
Db 241 YLSNSRIPNPIEGMSLSPDQTNVQFNGRCTIDGQPLGTTTPVSVSOLCKFRGRTISGQR 300
QY 301 VLNLTELDGSPFMAFAAPAGPDLGSCDWHIEMSKIPNSSTONNPIVTVNSVKNSQOF 360
Db 301 VLNLTELDGSPFMAFAAPAGPDLGSCDWHIEMSKIPNSSTONNPIVTVNSVKNSQOF 360
QY 361 VPHLSITLDDNVSSGGDIYGTIQTWSPPSDGGANTFWKIPDYGSSLAESAQLAPAY 420
Db 361 VPHLSITLDDNVSSGGDIYGTIQTWSPPSDGGANTFWKIPDYGSSLAESAQLAPAY 420
QY 421 PGFNEVIVYFMAIPGPNQSGSNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPT 480
Db 421 PGFNEVIVYFMAIPGPNQSGSNLVPCLLPQEIYTHFISEQAPIQGEAALLHYVDPT 480
QY 481 NRNLGEFKLYPGGYLTCVPNSSTGQQLPLDGVFVFASWSRFFYQLKPVGTAGPARGRL 540
Db 481 NRNLGEFKLYPGGYLTCVPNSSTGQQLPLDGVFVFASWSRFFYQLKPVGTAGPARGRL 540
QY 541 GVR 544
Db 541 GVR 544

```

## RESULT 5

```

Q91EJ2  PRELIMINARY; PRT; 544 AA.
ID Q91EJ2
AC Q91EJ2;
DT 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)
DE Nucleocapsid protein.
OS Human calicivirus Hu/NLV/Queen's Arms/Leeds/92/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=159480;
RN [1]
RP SEQUENCE FROM N.A.
RA Clegg C.S., Chamberlain J., Green J., Brown D.W.G., Lewis D.;
RT "Analysis of diversity in the capsid and ORF3 proteins of some
RT Norwalk-like viruses circulating in the U.K.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.

```





```
OC Norovirus.
OX NCBI_TaxID=122920;
RN SEQUENCE FROM N.A.
RP STRAIN=HU/NLV/Valetta/95/Malta;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J.C., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL: AJ277616; CAB89097.1; -.
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
FT NON_TER 1
FT TER 544
SQ SEQUENCE 544 AA; 58415 MW; B6741846BD82E6D1 CRC64;

Query Match          97.7%; Score 2825; DB 12; Length 544;
Best Local Similarity 98.0%; Pred. No. 1.2e-219;
Matches 533; Conservative 4; Mismatches 7; Indels 0; Gaps 0;

QY 1 MMWASKDAPPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLIDPWIINF 60
   |||||
Db 1 MMWASKDAPPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLIDPWIINF 60
   |||||

QY 61 VQAPQGEFTISPNTPGDLVFDLQGLPHLNPFLSHLSQMYNGWGMNRRVRLVLAGNAFTA 120
   |||||
Db 61 VQAPQGEFTISPNTPGDLVFDLQGLPHLNPFLSHLSQMYNGWGMNRRVRLVLAGNAFTA 120
   |||||

QY 121 GKVIICCVPPGQSRSTLSIAQATLPHVIADVRTLDPEVEPLEDVRLVYHNDTQPTMR 180
   |||||
Db 121 GKVIICCVPPGQSRSTLSIAQATLPHVIADVRTLDPEVEPLEDVRLVYHNDTQPTMR 180
   |||||

QY 181 LLCMLYTLRTGASGGTDSFVAVAGRLTCPCGDFNLFVLPPTVEQKTRPFTVPNIPLK 240
   |||||
Db 181 LLCMLYTLRTGASGGTDSFVAVAGRLTCPCGDFNLFVLPPTVEQKTRPFTVPNIPLK 240
   |||||

QY 241 YLSNRPINPIEGMSLSPDQTQNVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSGOR 300
   |||||
Db 241 YLSNRPINPIEGMSLSPDQTQNVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSGOR 300
   |||||

QY 301 VLNLTELDSPPMAFAAPAGAPFDLGSCHWIEMSKIPNSSTQNNPIVNSVKNSQOF 360
   |||||
Db 301 VLNLTELDSPPMAFAAPAGAPFDLGSCHWIEMSKIPNSSTQNNPIVNSVKNSQOF 360
   |||||

QY 361 VPHLSITLDENVSSGGDYIGTIQWTSPPDSGGANTFWKIPDYGSSLAELASQLAPAVY 420
   |||||
Db 361 VPHLSITLDENVSSGGDYIGTIQWTSPPDSGGANTFWKIPDYGSSLAELASQLAPAVY 420
   |||||

QY 421 PGFNEVIYVFMASIPGNQSGPNLVPCLLPQEVITHFISEQAPIQGEAALLHYVDPT 480
   |||||
Db 421 PLVFNVEIYVFMACIRGNQCGSNLVPCLSPOEVITHFISEQAPIQGEAALLHYVDPT 480
   |||||

QY 481 NRNLGEFKLYPGGYLTCVPSNSSLTGQQLPLDGVVFASWSRFRYQLKPVGTAGARGRL 540
   |||||
Db 481 NRNLGEFKLYPGGYLTCVPSNSSLTGQQLPLDGVVFASWSRFRYQLKPVGTAGARGRL 540
   |||||

QY 541 GVRV 544
   |||||
Db 541 GVRV 544
   |||||

RESULT 8
Q8V7J5 PRELIMINARY; PRT; 531 AA.
AC Q8V7J5;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein (Fragment).
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;

SEQUENCE FROM N.A.
STRAIN=HU/NLV/Thistlehall/90/UK;
MEDLINE=20404883; PubMed=10949950;
RA Green J.C., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL: AJ277616; CAB89097.1; -.
DR InterPro: IPR004005; Calici_coat.
DR InterPro: IPR008975; Viral_cap_coat.
DR Pfam: PF00915; Calici_coat; 1.
FT NON_TER 1
FT TER 531
SQ SEQUENCE 531 AA; 56747 MW; 9688132FDDCFB88F CRC64;

Query Match          97.3%; Score 2812; DB 12; Length 531;
Best Local Similarity 99.2%; Pred. No. 1.3e-218;
Matches 527; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 8 ATPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLIDPWIINFPVQAPQGE 67
   |||||
Db 1 ATPSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLIDPWIINFPVQAPQGE 60
   |||||

QY 68 FTISPNTPGDLVFDLQGLPHLNPFLSHLSQMYNGWGMNRRVRLVLAGNAFTAAGVVICC 127
   |||||
Db 61 FTISPNTPGDLVFDLQGLPHLNPFLSHLSQMYNGWGMNRRVRLVLAGNAFTAAGVVICC 120
   |||||

QY 128 VPPGQSRSTLSIAQATLPHVIADVRTLDPEVEPLEDVRLVYHNDTQPTMRLLCMLYT 187
   |||||
Db 121 VPPGQSRSTLSIAQATLPHVIADVRTLDPEVEPLEDVRLVYHNDTQPTMRLLCMLYT 180
   |||||

QY 188 PLRTGASGGTDSFVAVAGRLTCPCGDFNLFVLPPTVEQKTRPFTVPNIPLKYLNSRI 247
   |||||
Db 181 PLRTGASGGTDSFVAVAGRLTCPCGDFNLFVLPPTVEQKTRPFTVPNIPLKYLNSRI 240
   |||||

QY 248 PNPIEGMSLSPDQTQNVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSGORVNLNLTBL 307
   |||||
Db 241 PNPIEGMSLSPDQTQNVQFQNGRCTIDGQPLGTTTPVSVSQLCKFRGRTSGORVNLNLTBL 300
   |||||

QY 308 DGSPMAFAAPAGAPFDLGSCHWIEMSKIPNSSTQNNPIVNSVKNSQOFVPHLSII 367
   |||||
Db 301 DGSPMAFAAPAGAPFDLGSCHWIEMSKIPNSSTQNNPIVNSVKNSQOFVPHLSII 360
   |||||

QY 368 TLDENVSSGGDYIGTIQWTSPPDSGGANTFWKIPDYGSSLAELASQLAPAVYPPGFNEV 427
   |||||
Db 361 TLDENVSSGGDYIGTIQWTSPPDSGGANTFWKIPDYGSSLAELASQLAPAVYPPGFNEV 420
   |||||

QY 428 IVYFMASIPGNQSGPNLVPCLLPQEVITHFISEQAPIQGEAALLHYVDPTNRNLGEF 487
   |||||
Db 421 IVYFMASIPGNQSGPNLVPCLLPQEVITHFISEQAPIQGEAALLHYVDPTNRNLGEF 480
   |||||

QY 488 KLYPGGYLTCVPSNSSLTGQQLPLDGVVFASWSRFRYQLKPVGTAGARG 538
   |||||
Db 481 KLYPGGYLTCVPSNSSLTGQQLPLDGVVFASWSRFRYQLKPVGTAGARG 531
   |||||

RESULT 9
Q9IV36 PRELIMINARY; PRT; 544 AA.
AC Q9IV36;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Thistlehall/90/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122925;
RN [1]
RP SEQUENCE FROM N.A.
STRAIN=HU/NLV/Thistlehall/90/UK;
```



```
RC STRAIN=SUZUGI;
RX MEDLINE=22192455; PubMed=12202225;
RA Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,
RA Gojobori T., Takeda N.;
RT "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
RT Viruses.";
RL Virology 299:225-239(2002).
DR EMBL; AB039774; BAC11813.1; -
DR InterPro; IPR004005; Calici_coat.
DR Pfam; PF00915; Calici_coat_1.
SQ SEQUENCE 543 AA; 58761 MW; E0DE7489A9C3498C CRC64;

Query Match      81.2%; Score 2347.5; DB 12; Length 543;
Best Local Similarity 79.5%; Pred. No. 4.5e-181;
Matches 434; Conservative 50; Mismatches 57; Indels 5; Gaps 3;

QY 1 MMASKDAPSDAGATGATGAGQLVPEVNTADPIPIIDPVAGSSTALATAGQVNLDPWLNFF 60
DB 1 MMASKDAPSDAGATGAGQLVPEVNTADPIPIIDPVAGSSTALATAGQVNLDPWLNFF 60
QY 61 VOAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSOMYNGWGMVRVVLGNAPTA 120
DB 61 VOAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSOMYNGWGMVRVVLGNAPTA 120
QY 121 GKVIICCVPPGQSRTLSIAQATLFPFHVIADVRTLDPEVPLEDVNRVLYHNNDTQPTMR 180
DB 121 GKVIICCVPPGQSRTLSIAQATLFPFHVIADVRTLDPEVPLEDVNRVLYHNNDTQPTMR 180
QY 181 LLCMLYTPRTGASGGTDSFVAGRVLTCPGDPNLFPLVPTVEQKTRPTVNPILK 240
DB 181 LLCMLYTPRTGASGGTDSFVAGRVLTCPGDPNLFPLVPTVEQKTRPTVNPILK 240
QY 241 YLSNSRIENPIEGMSLSPDQNTQVQFNGRCITIDGQPLGTTTPVSVOLCKEGRITSGOR 300
DB 241 YLSNSRIENPIEGMSLSPDQNTQVQFNGRCITIDGQPLGTTTPVSVOLCKEGRITSGOR 300
QY 301 VLNLTELDGSPFMAFAPAGFPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQOF 360
DB 301 VLNLTELDGSPFMAFAPAGFPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQOF 360
QY 361 VPHLSSITLDENVSSGGYIGTIQWTSPPSD--SGGANTNFWKIPDYGSLSAEASOLAPA 418
DB 361 VPHLSSITLDENVSSGGYIGTIQWTSPPSD--SGGANTNFWKIPDYGSLSAEASOLAPA 418
QY 419 VYPGGEIVIVFMASIPGPNQSGSNLYPCLLPQEIYTHFISEQAPIQGEAALLHYVDP 478
DB 419 VYPGGEIVIVFMASIPGPNQSGSNLYPCLLPQEIYTHFISEQAPIQGEAALLHYVDP 478
QY 479 DTRNRLGEFKLYPGGYLTCVPNSSSTGTPQPLDGVFVFSVSRFYQLKPVGTAGPARG 538
DB 479 DTRNRLGEFKLYPGGYLTCVPNSSSTGTPQPLDGVFVFSVSRFYQLKPVGTAGPARG 538
QY 539 RLGVRR 544
DB 539 RLGIIR 543

RESULT 12
Q918B9 PRELIMINARY; PRT; 543 AA.
AC Q918B9;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Appalachicola Bay/318/1995/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171838;
RN [1]
```

```
RP SEQUENCE FROM N.A.
RX STRAIN=Hu/NLV/Appalachicola Bay/318/1995/US;
RC MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RX STRAIN=Hu/NLV/Appalachicola Bay/318/1995/US;
RC MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RX STRAIN=Hu/NLV/Appalachicola Bay/318/1995/US;
RC MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RX STRAIN=Hu/NLV/Appalachicola Bay/318/1995/US;
RC STRAIN=Hu/NLV/Appalachicola Bay/318/1995/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Roger G.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414406; AAL12971.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat_1_coat.
SQ SEQUENCE 543 AA; 58362 MW; 2COA035E80A20EF5 CRC64;

Query Match      81.2%; Score 2346.5; DB 12; Length 543;
Best Local Similarity 78.9%; Pred. No. 5.4e-181;
Matches 430; Conservative 49; Mismatches 63; Indels 3; Gaps 2;

QY 1 MMASKDAPSDAGATGATGAGQLVPEVNTADPIPIIDPVAGSSTALATAGQVNLDPWLNFF 60
DB 1 MMASKDAPSDAGATGAGQLVPEVNTADPIPIIDPVAGSSTALATAGQVNLDPWLNFF 60
QY 61 VOAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSOMYNGWGMVRVVLGNAPTA 120
DB 61 VOAPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSOMYNGWGMVRVVLGNAPTA 120
QY 121 GKVIICCVPPGQSRTLSIAQATLFPFHVIADVRTLDPEVPLEDVNRVLYHNNDTQPTMR 180
DB 121 GKVIICCVPPGQSRTLSIAQATLFPFHVIADVRTLDPEVPLEDVNRVLYHNNDTQPTMR 180
QY 181 LLCMLYTPRTGASGGTDSFVAGRVLTCPGDPNLFPLVPTVEQKTRPTVNPILK 240
DB 181 LLCMLYTPRTGASGGTDSFVAGRVLTCPGDPNLFPLVPTVEQKTRPTVNPILK 240
QY 241 YLSNSRIENPIEGMSLSPDQNTQVQFNGRCITIDGQPLGTTTPVSVOLCKEGRITSGOR 300
DB 241 YLSNSRIENPIEGMSLSPDQNTQVQFNGRCITIDGQPLGTTTPVSVOLCKEGRITSGOR 300
QY 301 VLNLTELDGSPFMAFAPAGFPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQOF 360
DB 301 VLNLTELDGSPFMAFAPAGFPDLGSCDWHIEMSKIPNSSTQNNPIVTNSVKPNSQOF 360
QY 361 VPHLSSITLDENVSSGGYIGTIQWTSPPSD--SGGANTNFWKIPDYGSLSAEASOLAPA 419
DB 361 VPHLSSITLDENVSSGGYIGTIQWTSPPSD--SGGANTNFWKIPDYGSLSAEASOLAPA 419
QY 420 YPPGGEIVIVFMASIPGPNQSGSNLYPCLLPQEIYTHFISEQAPIQGEAALLHYVDP 479
DB 420 YPPGGEIVIVFMASIPGPNQSGSNLYPCLLPQEIYTHFISEQAPIQGEAALLHYVDP 479
QY 421 YPPGGEIVIVFMASIPGPNQSGSNLYPCLLPQEIYTHFISEQAPIQGEAALLHYVDP 480
```

```

QY 480 TNRNLGEFKLYPGGKLTCPVNSSTGCPQQLPLDGVFVFSWVSRFYQLKPVGTAGPARGR 539
DB 481 TNRNLGEFKLYPGEFVTCMPN--GTGPOQLPLNGVVFVFSWVSRFYQLKPVGTASSAKGR 538
QY 540 LGVRR 544
DB 539 LGVRR 543

RESULT 13
Q8JW44 PRELIMINARY; PRT; 546 AA.
AC Q8JW44;
DT 01-OCT-2002 (TReMBLrel. 22, Created)
DT 01-OCT-2002 (TReMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein.
GN CAPSID.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Funabashi258;
RA Katayama K., Takeda N., Natori K.;
RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=Funabashi258;
RA Natori K., Takeda N.;
RL "Genetic and antigenic relationship among Norwalk-like viruses.";
RT Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB078335; BAC05516.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 546 AA; 58834 MW; 1965F054E2C481E6 CRC64;

Query Match 77.2%; Score 2231.5; DB 12; Length 546;
Best Local Similarity 74.7%; Pred. No. 1.1e-171;
Matches 407; Conservative 59; Mismatches 78; Indels 1; Gaps 1;

QY 1 MMASKDAPQADGATGAGLVPEVNTADPIPDPVAGSSTALATAGVNLIDPWIIINNF 60
DB 1 MMASKDAPQADGATGAGLVPEVNTADPLPMEPVAGPTTAVATAGVNMIDPWVNNF 60
QY 61 VQAPQGEFTTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGVGNMVRVVLGNAFTA 120
DB 61 VQSPQGEFTTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMVRVILLAGNAFSA 120
QY 121 GKVIICCVPPGQSRRTLSIAQATLFPVHVIADVRTLDPEVEPLEDVRNVLVYHNDTQPTMR 180
DB 121 GKIIIVCCVPPGFTSSSLTIAQATLFPVHVIADVRTLEPIEMPLEDVRNVLVYHNDTQPTMR 180
QY 181 LLCMLYTLRTGGASGGTDSFVAVGRVLTCPGDPNFELFVPTTVEQKTRPFTVNIPLK 240
DB 181 LVCMLYTLRTGGSGNSDSFVAVGRVLTAPSSDFSFLFVPTTIEQKTRPFTVNIPLQ 240
QY 241 YLSNSRIINPIEGMSLSPDQNTQVQFNGRCLTIDGQPLGTTTPVSVSLCKFRGRITTSQR 300
DB 241 TLSNSRFPFLIOGMILSPDASQVQVQFNGRCLIDGQLLGTTPATSGQLFRVRGKINQAR 300
QY 301 VLNLTELDGSPFMAFAAPAGFPDLGSCDWHIEMSKIPINSSTQNNPIVTVNSVKPNSQOF 360
DB 301 TLNLTEVDGKPFMAFDSFAPVGFDFGKCDWHMRVSKTPNNTSSGDPMSRSVSVQTNVQGF 360
QY 361 VPHLSITLDENVIS--SGDYIGTIQWTSPPDSGGANTNFWKIPDYGSSLAESQLAPAV 419
DB 361 VPHLSIQDFEVNFHTPGDYIGTIEWISQPSPTPLGTDINLWEIPDYGSSLSQANLAPPV 420
QY 420 YPPGNEVTVYFWASIPGPNQSGSNLVPCLLPQBYITHFVSEQAPITQGEAALLHYVDPD 479
DB 421 FPPGFEALYFVFSAPFGPNNSAFNDVFCLLPQBYITHFVSEQAPITQGEAALLHYVDPD 480

```

```

DB 421 FPPGFEALYFVFSAPFGPNNSAFNDVFCLLPQBYITHFVSEQAPITQGEAALLHYVDPD 480
QY 480 TNRNLGEFKLYPGGKLTCPVNSSTGCPQQLPLDGVFVFSWVSRFYQLKPVGTAGPARGR 539
DB 481 TNRNLGEFKLYPGEFVTCMPN--GTGPOQLPLNGVVFVFSWVSRFYQLKPVGTASTARS 540
QY 540 LGVRR 544
DB 541 LGVRR 545

RESULT 14
Q9IV47 PRELIMINARY; PRT; 546 AA.
AC Q9IV47;
DT 01-OCT-2000 (TReMBLrel. 15, Created)
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NUV/Whiterose/96/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122914;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NUV/Whiterose/96/UK;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236 (2000).
DR EMBL; AJ277610; CAB89091.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 546 AA; 58737 MW; 501F9D2BE2A77B51 CRC64;

Query Match 77.0%; Score 2226.5; DB 12; Length 546;
Best Local Similarity 74.5%; Pred. No. 2.7e-171;
Matches 406; Conservative 61; Mismatches 77; Indels 1; Gaps 1;

QY 1 MMASKDAPQADGATGAGLVPEVNTADPIPDPVAGSSTALATAGVNLIDPWIIINNF 60
DB 1 MMASKGAPQADGAGGAGLVPEVNTADPLPMEPVAGPTTAVATAGVNMIDPWVNNF 60
QY 61 VQAPQGEFTTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGVGNMVRVVLGNAFTA 120
DB 61 VQSPQGEFTTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMVRVILLAGNAFSA 120
QY 121 GKVIICCVPPGQSRRTLSIAQATLFPVHVIADVRTLDPEVEPLEDVRNVLVYHNDTQPTMR 180
DB 121 GKIIIVCCVPPGFTSSSLTIAQATLFPVHVIADVRTLEPIEMPLEDVRNVLVYHNDTQPTMR 180
QY 181 LLCMLYTLRTGGASGGTDSFVAVGRVLTCPGDPNFELFVPTTVEQKTRPFTVNIPLK 240
DB 181 LVCMLYTLRTGGSGNSDSFVAVGRVLTAPSSDFSFLFVPTTIEQKTRPFTVNIPLQ 240
QY 241 YLSNSRIINPIEGMSLSPDQNTQVQFNGRCLTIDGQPLGTTTPVSVSLCKFRGRITTSQR 300
DB 241 TLSNSRFPFLIOGMILSPDASQVQVQFNGRCLIDGQLLGTTPATSGQLFRVRGKINQAR 300
QY 301 VLNLTELDGSPFMAFAAPAGFPDLGSCDWHIEMSKIPINSSTQNNPIVTVNSVKPNSQOF 360
DB 301 TLNLTEVDGKPFMAFDSFAPVGFDFGKCDWHMRVSKTPNNTSSGDPMSRSVSVQTNVQGF 360
QY 361 VPHLSITLDENVIS--SGDYIGTIQWTSPPDSGGANTNFWKIPDYGSSLAESQLAPAV 419
DB 361 VPHLSIQDFEVNFHTPGDYIGTIEWISQPSPTPLGTDINLWEIPDYGSSLSQANLAPPV 420
QY 420 YPPGNEVTVYFWASIPGPNQSGSNLVPCLLPQBYITHFVSEQAPITQGEAALLHYVDPD 479
DB 421 FPPGFEALYFVFSAPFGPNNSAFNDVFCLLPQBYITHFVSEQAPITQGEAALLHYVDPD 480

```

QY 480 TNRNLGEFKLYPGGYLTCVFNSSSTGPOQLPLDGVFVFASWVSRRFYQLKPVGTAGPARGR 539  
Db 481 TNRNLGEFKLYPGGYLTCVFNSSSTGPOQLPLDGVFVFASWVSRRFYQLKPVGTAGPARGR 540  
QY 540 LGVRR 544  
Db 541 LGVRR 545

RESULT 15  
Q8VA27 PRELIMINARY; PRT; 544 AA.  
AC Q8VA27;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus NLV/C59/99.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=177069;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Parkas T., Wilton N., Altaye M., Davies D., Chapman F., Thornton S.A.,  
Jiang X.;  
RT "Homologous vs. heterologous immune responses to Norwalk-like viruses  
among crew members following acute gastroenteritis outbreaks on two US  
Navy vessels";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF435807; AAL31551.1; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 544 AA; 58502 MW; D20034269B926E35 CRC64;

Query Match 76.7%; Score 2218.5; DB 12; Length 544;  
Best Local Similarity 74.4%; Pred. No. 1.2e-170;  
Matches 404; Conservative 60; Mismatches 78; Indels 1; Gaps 1;  
QY 3 MASKATPSADGATGAGQLVPEVTADPIPIPVAGSSTALATAGVNLIDPWIIINNFVQ 62  
Db 1 MASKAPQADGASGAGQLVPEVTADPIPIPVAGSSTALATAGVNLIDPWIIINNFVQ 60  
QY 63 APQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMRVVRVVLGNAFTAGK 122  
Db 61 SPQGEFTISPNNTPGDVLFDLQGLPHLNPFLSHLSQMYNGWGNMRVVRVVLGNAFTAGK 120  
QY 123 VIICVPPGFGSRTLSIAQATLFPVHIVADVRLDPVEVPLEDVRNVLHNNDTQPTMRL 182  
Db 121 IIVCCVPPGFTSSLTIAQATLFPVHIVADVRLDPVEVPLEDVRNVLHNNDTQPTMRL 180  
QY 183 CMLYPLRTGGASGGTDSFVAVAGVLTCPGDPNFLFLVPPTVEOKTRPFTVPIPLKYL 242  
Db 181 CMLYPLRTGGGSGSDSFVAVAGVLTCPGDPNFLFLVPPTVEOKTRPFTVPIPLKYL 240  
QY 243 SNRIPNPIEGMSLSPTQNVQFQNGRCTIDQPLGTTTPSVSOLCKFRGRTSGQRYL 302  
Db 241 SNRIPNPIEGMSLSPTQNVQFQNGRCTIDQPLGTTTPSVSOLCKFRGRTSGQRYL 300  
QY 303 NLTLDGSPFMAFAPAPAGFPDLAGSCDWHIEMSKI PNSTQNNPIVTVNSVKNPQGFVP 362  
Db 301 NLTLDGSPFMAFAPAPAGFPDLAGSCDWHIEMSKI PNSTQNNPIVTVNSVKNPQGFVP 360  
QY 363 HLLSITLDENVIS-SGGDYITGTTQWTSPPSDSGGANTFWKI PDYGSLSAEASQLAPAVYP 421  
Db 361 HLLSITLDENVIS-SGGDYITGTTQWTSPPSDSGGANTFWKI PDYGSLSAEASQLAPAVYP 420  
QY 422 PGFNEVIVFMASIPGNQSGSNLVPCLLPQEYITHFISEQAPIQGEAALLHYVDPDTN 481  
Db 421 PGFNEVIVFMASIPGNQSGSNLVPCLLPQEYITHFISEQAPIQGEAALLHYVDPDTN 480  
QY 482 RNLGEFKLYPGGYLTCVFNSSSTGPOQLPLDGVFVFASWVSRRFYQLKPVGTAGPARGR 541

Db 481 RNLGEFKLYPGGYLTCVFNSSSTGPOQLPLDGVFVFASWVSRRFYQLKPVGTAGPARGR 540  
QY 542 VRR 544  
Db 541 VRR 543

Search completed: June 1, 2004, 13:53:13  
Job time : 32.6725 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 46.3034 Seconds  
(without alignments)  
3289.030 Million cell updates/sec

Title: US-09-926-799-5  
Perfect score: 2886  
Sequence: 1 MKMASNDANPSDGTANLVP.....VNQFTYLPAMGNGAGRRAL 539

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2886	100.0	539	AAB49704	Aab49704 Small rou
2	1947.5	67.5	548	AAB49705	Aab49705 Small rou
3	1919	66.5	535	AAB49707	Aab49707 Small rou
4	1907.5	66.1	548	Aau91272	Aau91272 Norwalk v
5	1872.5	64.9	540	AAB49706	Aab49706 Small rou
6	1834.5	63.6	550	AAB49709	Aab49709 Small rou
7	1821.5	63.1	542	AAB49708	Aab49708 Small rou
8	1569	54.4	541	AAB49710	Aab49710 Small rou
9	1161.5	40.2	530	AAR50972	Aar50972 Norwalk v
10	1161.5	40.2	530	ADC72176	Adc72176 Norwalk v
11	1157.5	40.1	545	AAB49700	Aab49700 Small rou
12	1151.5	39.9	544	AAB49703	Aab49703 Small rou
13	1148.5	39.8	530	AAB49701	Aab49701 Small rou
14	1106.5	38.3	530	AAR57091	Aar57091 Small rou
15	1105	38.3	546	AAB49702	Aab49702 Small rou
16	322	11.2	669	AAB67461	Aab67461 Amino aci
17	321	11.1	579	AAB08143	Aab08143 RHDV caps
18	303	10.5	623	AAB47044	Aab47044 Feline ca
19	300.5	10.4	547	AAM50107	Aam50107 Feline ca
20	300.5	10.4	671	AAM50108	Aam50108 Feline ca
21	300	10.4	623	AAB47043	Aab47043 Feline ca
22	298	10.3	622	AAB47045	Aab47045 Feline ca
23	297.5	10.3	668	AAB67462	Aab67462 Amino aci
24	293	10.2	668	AAR10686	Aar10686 Feline ca
25	293	10.2	668	AAB04304	Aab04304 Feline ca

26	197	6.8	40	5	Aau91274	Aau91274 Norwalk v
27	184	6.4	40	5	Aau91273	Aau91273 Norwalk v
28	135	4.7	6310	6	ABU39869	Abu39869 Protein e
29	130.5	4.5	934	1	AAP20016	Aap20016 Sequence
30	126.5	4.4	1147	5	ABB76724	Abb76724 Foot and
31	125.5	4.3	31267	6	ABG74786	Abg74786 Human RGS
32	122	4.2	26926	4	Aau05396	Aau05396 Human tit
33	121	4.2	2406	4	ABB64198	Abb64198 Drosophil
34	117	4.1	998	4	ABU53140	Abu53140 Human tse
35	117	4.1	1024	7	ADC11296	Adc11296 Human nov
36	116	4.0	5701	4	ABB36684	Abb36684 Peptide #
37	116	4.0	5701	4	ABB22021	Abb22021 Protein #
38	116	4.0	5701	4	AAM69843	Aam69843 Human bon
39	116	4.0	5701	4	AAM57448	Aam57448 Human bra
40	116	4.0	5701	4	ABG51536	Abg51536 Human liv
41	116	4.0	5701	4	AAM05328	Aam05328 Peptide #
42	114.5	4.0	2975	6	ABM68759	Abm68759 Photorhab
43	112.5	3.9	710	6	ADA54747	Ada54747 Human pro
44	112	3.9	1344	4	AAG93134	Aag93134 C glutami
45	111.5	3.9	776	2	AAM55105	Aam55105 Streptoco

ALIGNMENTS

RESULT 1  
AAB49704  
ID AAB49704 standard; protein; 539 AA.

AC AAB49704;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 5.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX  
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.

Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
WPI: 2001-080848/09.  
DR N-PSDB; AAF29145.  
PT Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.  
XX Claim 1; Page 50-52; 84pp; Japanese.  
PS This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks  
SQ Sequence 539 AA;  
Query Match 100.0%; Score 2886; DB 4; Length 539;  
Best Local Similarity 100.0%; Pred. No. 1.1e-251;

```

Matches 539; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKMASNDANPSDGSSTANLVPVNEVMALPVPVGAIAAPVAGQONVDPWIRNNFVOAP 60
DB 1 MKMASNDANPSDGSSTANLVPVNEVMALPVPVGAIAAPVAGQONVDPWIRNNFVOAP 60
QY 61 GGEFTVSPRNAPGEILWSAPLQPDLPNLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNAPGEILWSAPLQPDLPNLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
QY 121 FFAVPPNPFTEGLSPSQVTMFPPIIIVDVRQLEPVLPLPDPVNNFYHYNQSDSTIKLIA 180
DB 121 FFAVPPNPFTEGLSPSQVTMFPPIIIVDVRQLEPVLPLPDPVNNFYHYNQSDSTIKLIA 180
QY 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDFDFIFLVPPTVESRTKPTFVPIILTVBEEMNS 240
DB 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDFDFIFLVPPTVESRTKPTFVPIILTVBEEMNS 240
QY 241 RFPPIPLEKLYTGPSSAFVVPQNGRCTTQDGVLLGTTQLSAVNICTFRGDTVHIAGSHDYT 300
DB 241 RFPPIPLEKLYTGPSSAFVVPQNGRCTTQDGVLLGTTQLSAVNICTFRGDTVHIAGSHDYT 300
QY 301 MNLASQWNSNYDPTETBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVFTF 360
DB 301 MNLASQWNSNYDPTETBEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSHVFTF 360
QY 361 KLGSVQYTTDTNNDFTQNTKFTPVGVIQDGNHQNHPQWVLPNYSGRTHNVHLAPA 420
DB 361 KLGSVQYTTDTNNDFTQNTKFTPVGVIQDGNHQNHPQWVLPNYSGRTHNVHLAPA 420
QY 421 VAPTPEGEQLLFFRSTMPGCSGYPNMNLDCLLPOEWWQHFCEAAQAQSDVALLRFVNP 480
DB 421 VAPTPEGEQLLFFRSTMPGCSGYPNMNLDCLLPOEWWQHFCEAAQAQSDVALLRFVNP 480
QY 481 TGRVLECKLHKSQYVTVTAHTGPHDLVIPPNGYFRFDSWVNFYTLAPMNGAGRRAL 539
DB 481 TGRVLECKLHKSQYVTVTAHTGPHDLVIPPNGYFRFDSWVNFYTLAPMNGAGRRAL 539

RESULT 3
AAB49705
ID AAB49705 standard; protein; 548 AA.
XX
AC AAB49705;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 6.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX WO200079280-A1.
XX
XX 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
XX
XX N-PSDB; AAF29146.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX
XX strains for investigation of food poisoning outbreaks, contains
XX
XX antibodies.
XX

```

---

```

PS Claim 1; Page 52-54; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 548 AA;
QY Query Match 67.5%; Score 1947.5; DB 4; Length 548;
DB Best Local Similarity 66.2%; Pred. No. 8.2e-167;
QY Matches 365; Conservative 73; Mismatches 94; Indels 19; Gaps 4;
DB 1 MKMASNDANPSDGSSTANLVPVNEVMALPVPVGAIAAPVAGQONVDPWIRNNFVOAP 60
DB 1 MKMASNDANPSDGSSTANLVPVNEVMALPVPVGAIAAPVAGQONVDPWIRNNFVOAP 60
QY 61 GGEFTVSPRNAPGEILWSAPLQPDLPNLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNAPGEILWSAPLQPDLPNLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
QY 121 FFAVPPNPFTEGLSPSQVTMFPPIIIVDVRQLEPVLPLPDPVNNFYHYNQSDSTIKLIA 180
DB 121 FFAVPPNPFTEGLSPSQVTMFPPIIIVDVRQLEPVLPLPDPVNNFYHYNQSDSTIKLIA 180
QY 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDFDFIFLVPPTVESRTKPTFVPIILTVBEEMNS 240
DB 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDFDFIFLVPPTVESRTKPTFVPIILTVBEEMNS 240
QY 241 RFPPIPLEKLYTGPSSAFVVPQNGRCTTQDGVLLGTTQLSAVNICTFRGDTVHIAG----- 295
DB 241 RFPPIPLEKLYTGPSSAFVVPQNGRCTTQDGVLLGTTQLSAVNICTFRGDTVHIAG----- 295
QY 296 -----SHDYTMNLASQWNSNYDPTETBEIPAPLGTDPFVGKIQGMLTQTTREDGSTR 346
DB 296 -----SHDYTMNLASQWNSNYDPTETBEIPAPLGTDPFVGKIQGMLTQTTREDGSTR 346
QY 301 ADTATPRLFNYYHVVQDNLNGTTPYDPAEDIPGPGTDFRGKVFQVSGASQRL--DSTRA 359
DB 301 ADTATPRLFNYYHVVQDNLNGTTPYDPAEDIPGPGTDFRGKVFQVSGASQRL--DSTRA 359
QY 347 HKATVSTGSHVFTFVKLGSVQYTTDTNNDFTQNTKFTPVGVIQDGNHQNHPQWVLPN 406
DB 347 HKATVSTGSHVFTFVKLGSVQYTTDTNNDFTQNTKFTPVGVIQDGNHQNHPQWVLPN 406
QY 360 HEAKYDVTAGRTFKLGSLEISTD--SDDFDQNTKFTPVGI---GYDNEAEFQWLSLP 415
DB 360 HEAKYDVTAGRTFKLGSLEISTD--SDDFDQNTKFTPVGI---GYDNEAEFQWLSLP 415
QY 407 YSGRTGHNHVLAPAVAPTFFGEQLLFFRSTMPGCSGYPNMNLDCLLPOEWWQHFCEAAAP 466
DB 407 YSGRTGHNHVLAPAVAPTFFGEQLLFFRSTMPGCSGYPNMNLDCLLPOEWWQHFCEAAAP 466
QY 416 YSGQFTHNMNLAPAVAPNPFGEQLLFFRSTMPGCSGYPNMNLDCLLPOEWWQHFCEASAP 475
DB 416 YSGQFTHNMNLAPAVAPNPFGEQLLFFRSTMPGCSGYPNMNLDCLLPOEWWQHFCEASAP 475
QY 467 AQSDVALLRFVNPDTGRVLFECLEKHSQYVTVTAHTGPHDLVIPPNGYFRFDSWVNFYTL 526
DB 467 AQSDVALLRFVNPDTGRVLFECLEKHSQYVTVTAHTGPHDLVIPPNGYFRFDSWVNFYTL 526
QY 476 AQTOVALVRYVNPDTGKVLFEAKLHKLGMFTIANGSDSPITVPNGYFRFDSWVNFYTL 535
DB 476 AQTOVALVRYVNPDTGKVLFEAKLHKLGMFTIANGSDSPITVPNGYFRFDSWVNFYTL 535
QY 527 APMGNGAGRRR 537
DB 536 APMGTGNGRRR 546

RESULT 3
AAB49707
ID AAB49707 standard; protein; 535 AA.
XX
XX AAB49707;
XX
XX AC AAB49707;
XX
XX DT 04-APR-2001 (first entry)
XX
XX DE Small round structured virus protein SEQ ID 8.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX WO200079280-A1.
XX

```



```
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000MO-JF004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI: 2001-080848/09.
XX N-PSDB; AAF29148.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
XX Claim 1; Page 57-59; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 535 AA;
SQ
Query Match 66.5%; Score 1919; DB 4; Length 535;
Best Local Similarity 65.4%; Pred. No. 3e-164;
Matches 353; Conservative 80; Mismatches 97; Indels 10; Gaps 5;
QY 1 MKMASNDANPSGSGTANLVPEVNEVMALPEVVGAAIAAPVAGQONVIDPMIRNMFVQAP 60
DB 1 MKMASNDANPSGSGTANLVPEVNEVMALPEVVGAAIAAPVAGQONVIDPMIRNMFVQAP 60
QY 61 GGEFTVSPRNPARGELIWSAPLGDPLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
DB 61 NGEFTVSPRNPARGELIWSAPLGDPLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
QY 121 FAAPVPPNPTGSLSPQVMTFPHIIVDVRLQEPVLIPLDVVRNNFYHYNQSDSTIKLIA 180
DB 121 FAAPVPPNPTGSLSPQVMTFPHIIVDVRLQEPVLIPLDVVRNNFYHYNQSDSTIKLIA 180
QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFELVPPPTVESRTKPFVPIILTVEMSGS 240
DB 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFELVPPPTVESRTKPFVPIILTVEMSGS 240
QY 241 RPPIPLEKLYTGPSAFVQVQNGRCCTTGDVLLGTTQLSAVNICITFRGDVTH- IAG-SHD 298
DB 241 RFPVPIDELITYSPNLSLVQVQNGRCALDGELOGTQLLPTALCSFRGEINQKVSQENHV 300
QY 299 YTMNLASQNSWYDPTTEIPAPLGTPDFVGKIQGMLTQTTREDGSTRAHKATVSTGSHV 358
DB 301 WNMQVTNINGTPFDGDPVAPLGTDFSGKLGVLQSQRD-HDNACRSHDAVIATNSAKF 359
QY 359 TPKLASVQYTTTNDNFQTKTQNTKTPVGVQV- DGNHQNHPQOQWLVNYSRGTGHVHL 417
DB 360 TPKLGAIQGTWEEDDVHINQPTKPTVGLFENEGFN-----QWTLNYSGALTINMGL 413
QY 418 APAAVPTFPEQLLFRSMTMPCSGYPNNMLDCLLPQEWVQHFCEBAAQAQSDVALLRFV 477
DB 414 APVAPTFFGEQILFRSHIPLKGVADPVIQCLLPQEWIQLHYESAPSQSDVALIRFT 473
QY 478 NPDTCRVLFECKLHSGYVTVARTGPHDLVIPNGYFRFDSWVNOFYTLAPMNGAGRRR 537
DB 474 NPDTCRVLFEAKLHRSYITVANTGSRPIVVPVANGYFRFDTWVNOFYSLAPMGTGNGRRR 533
RESULT 4
AAU91272
```

```
ID AAU91272 standard; protein; 548 AA.
XX
AC AAU91272;
XX
DT 18-JUN-2002 (first entry)
XX
DE Norwalk virus associated polynucleotide #1.
XX
XX Norwalk virus; monoclonal antibody; geno group I; geno group II;
KW immunological detection; food; viral infection.
XX
XX Norwalk virus.
XX
XX JP2002020399-A.
XX
XX 23-JAN-2002.
XX
XX 10-JUL-2000; 2000JP-00208151.
XX
XX 10-JUL-2000; 2000JP-00208151.
XX (OSAP ) OSAKA PREFECTURE.
PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.
XX (IATR ) IATRON LAB INC.
XX
XX WPI: 2002-287412/33.
XX
XX A monoclonal antibody useful in the immunological detection and diagnosis
PT of Norwalk virus infection.
XX
XX Disclosure; Page 12-13; 24pp; Japanese.
XX
XX The invention describes a monoclonal antibody recognising Norwalk virus,
CC a capsid protein of Norwalk virus, or a common antigen epitope on the
CC capsid protein molecule of geno group I and geno group II. The antibody
CC is useful for immunological detection and quantitative analysis of
CC Norwalk virus in foods and the serum of infected patients. This sequence
CC represents a Norwalk virus associated protein described in the invention
XX
XX Sequence 548 AA;
Query Match 66.1%; Score 1907.5; DB 5; Length 548;
Best Local Similarity 65.2%; Pred. No. 3.4e-163;
Matches 358; Conservative 75; Mismatches 97; Indels 19; Gaps 4;
QY 1 MKMASNDANPSGSGTANLVPEVNEVMALPEVVGAAIAAPVAGQONVIDPMIRNMFVQAP 60
DB 1 MKMASNDANPSGSGTANLVPEVNEVMALPEVVGAAIAAPVAGQONVIDPMIRNMFVQAP 60
QY 61 GGEFTVSPRNPARGELIWSAPLGDPLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
DB 61 GGEFTVSPRNPARGELIWSAPLGDPLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAGKII 120
QY 121 FAAPVPPNPTGSLSPQVMTFPHIIVDVRLQEPVLIPLDVVRNNFYHYNQSDSTIKLIA 180
DB 121 FAAPVPPNPTGSLSPQVMTFPHIIVDVRLQEPVLIPLDVVRNNFYHYNQSDSTIKLIA 180
QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFELVPPPTVESRTKPFVPIILTVEMSGS 240
DB 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFELVPPPTVESRTKPFVPIILTVEMSGS 240
QY 241 RPPIPLEKLYTGPSAFVQVQNGRCCTTGDVLLGTTQLSAVNICITFRGDVTHIAG----- 295
DB 241 RFPVPIESLHTSPENIVVQVQNGRCVTLGELMGTTQLLPSRICAFRGVLTSTSRASDQ 300
QY 296 -----SHDYTMNLASQNSWYDPTTEIPAPLGTPDFVGKIQGMLTQTTREDGSTRA 346
DB 301 ADIATPRLFNYWYVQLDNLNGTYPDAEDIPGLTDFDFRGKVFVASQ-RNPDSITTRA 359
QY 347 HKATVSTGSHVFTPKLGSVQYTTTNDNFQTKTQNTKTPVGVQVQDGNHQNHPQOQWLVN 406
DB 360 HEAKVDITAGRTPKLGSLEISTE-SSDFQDQNPQTRFTPVGI---GVONEADFOQWSLDP 415
```



CC spread of infection and to examine the epidemiology of outbreaks

XX SQ Sequence 550 AA;

Query Match 63.6%; Score 1834.5; DB 4; Length 550;  
Best Local Similarity 62.5%; Pred. No. 1.4e-156;  
Matches 348; Conservative 85; Mismatches 97; Indels 27; Gaps 7;  
QY 1 MKMASNDANPSGSTANLVPEVNEVMALEPVVGAIAAPVAGQNVQVDPWIRNPFVQAP 60  
DB 1 MKMASNDAAAPSDGAAANLVPEANDEVMALEPVVGAIAAPVVGQNVQVDPWIRNPFVQAP 60  
QY 61 GSEFTVSPNAPGELLWSAPLGNLPSLHARMYNGYAGGFEVQVLLAGNAFTAGKII 120  
DB 61 QGEFTVSPNAPGELLWSAPLGNLPSLHARMYNGYAGGFEVQVLLAGNAFTAGKII 120  
QY 121 FFAVPPNFTPEGLSPQVTFPHIIVDVRQLEPVLPLPDVNRNFFHYNQSDNFKILIA 180  
DB 121 FFAVPPHPPVENISAAQITMCPHVIVDVRQLEPVLPLPDVNRNFFHYNQSDNFKILIA 180  
QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFIPLVPTVESRTKPTVPILTVEEMNS 240  
DB 181 MLYTPLRA-NSGEDVFTVSCRVLTRPADDFEFLVPTVESRTKPTVPILTVEEMNS 239  
QY 241 RPPIDPLEKLYTGPSSAFVQVQNGRCTTGVLLGTQLSNAVNICTFRG----- 288  
DB 240 RPPAAIDMLYTDPNESIVVQVQNGRCTLDGTLQGTQLVPTQICAFRGTLISQTARAADS 299  
QY 289 -DVTHIAGSHDYTMNLASONWSNYDPTBEIPAPLGTDPDFVGIQGLTQT---TREDG- 342  
DB 300 TDSPOARHPLHVQVKNLDGTYDPTDIPAVLGAIDFGVGVASORDVSGQOQGH 359  
QY 343 -STRAHKAIVTSGSVHFTPKLGSVQVTTDTNNDFOTGQNTKPTPVGVIQDGNHQNHPQQ 401  
DB 360 YATRAHEAHDITDPKYAPKLTILKSG-SDENTNQPIRTPVGM---GDNW---RQ 412  
QY 402 WVLPNYSGRTHNVHLAPAVATFPGEQLLFRSTMPGCGSPNNMLDCLLPQEWVQHFC 461  
DB 413 WELPDYSGRLTNMLNAPAVSPSPGERILFRSIVPSAGGYGSGYIDCLIPQEWVQHFY 472  
QY 462 QEAAPQAQSDVALLRFVNPDTGRVLFECCLKHSGYVTVAGTGHDPHDIVIPNGYFRFDSWN 521  
DB 473 QEAAPQAQSDVALLRFVNPDTGRVLFECCLKHSGYVTVAGTGHDPHDIVIPNGYFRFDSWN 521  
QY 522 QFYTLAPMGNGAGRRRA 538  
DB 533 QFYTLAPMGSGGQRRRA 549

RESULT 7  
AAB49708  
ID AAB49708 standard; protein; 542 AA.

XX AC AAB49708;

XX DT 04-APR-2001 (first entry)

XX DE Small round structured virus protein SEQ ID 9.

XX KW Small round structured virus; SRSV; food poisoning.

XX OS Small round structured virus.

XX PN WO200079280-A1.

XX PD 28-DEC-2000.

XX PF 22-JUN-2000; 2000WO-JP004095.

XX PR 22-JUN-1999; 99JP-00175928.

XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

XX (DENK-) DENKA SEIKEN KK.

XX

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX WPI; 2001-080848/09.

DR N-PSDB; AAF29149.

XX PT Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
antibodies.

XX PS Claim 1; Page 59-61; 84pp; Japanese.

XX CC This invention relates to a kit for the detection and typing of small  
CC round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks

XX SQ Sequence 542 AA;

Query Match 63.1%; Score 1821.5; DB 4; Length 542;

Best Local Similarity 63.7%; Pred. No. 2e-155;

Matches 347; Conservative 74; Mismatches 111; Indels 13; Gaps 6;

QY 1 MKMASNDANPSGSTANLVPEVNEVMALEPVVGAIAAPVAGQNVQVDPWIRNPFVQAP 60

DB 1 MKMASNDAAAPSDGAAANLVPEGINETMPLEPVAGASIAAPVAGQNIIDPWIRNPFVQAP 60

QY 61 GGEFTVSPNAPGELLWSAPLGNLPSLHARMYNGYAGGFEVQVLLAGNAFTAGKII 120

DB 61 NGEFTVSPNAPGELLWSAPLGNLPSLHARMYNGYAGGFEVQVLLAGNAFTAGKIL 120

QY 121 FFAVPPNFTPEGLSPQVTFPHIIVDVRQLEPVLPLPDVNRNFFHYNQSDNFKILIA 180

DB 121 FFAVPPNFTPEGLSPQVTFPHIIVDVRQLEPVLPLPDVNRNFFHYNQSDNFKILIA 180

QY 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFIPLVPTVESRTKPTVPILTVEEMNS 240

DB 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFIPLVPTVESRTKPTVPILTVEEMNS 240

QY 241 RPPIDPLEKLYTGPSSAFVQVQNGRCTTGVLLGTQLSNAVNICTFRGDVTIAGSH--D 298

DB 241 RPPIDPLEKLYTGPSSAFVQVQNGRCTTGVLLGTQLSNAVNICTFRGDVTIAGSH--D 298

QY 299 YTMNLASONWSNYDPTBEIPAPLGTDPDFVGIQGLTQTTRDGSSTRAHKAT---VSTGS 355

DB 301 NLLQTYPNGASYDPTBEIPAPLGTDPDFVGIQGLTQTTRDGSSTRAHKAT---VSTGS 355

QY 356 VHFPTPKLGSVQVTTDTNNDFOTGQNTKPTPVGVIQDGNHQNHP-QQWVLPNYSGRTHN 414

DB 361 GKFTPKIGSIGLHSITEH-VHPNQSRFTPVGAVD---ENTPQQWVLPNYSGRTHN 415

QY 415 VHLAPAVAPTFFGQQLLFRSTMPGCGSGYPNN--LDCLLPQEWVQHFCQAPASQVDA 472

DB 416 TNLAPAVAPTFFGQQLLFRSTMPGCGSGYPNN--LDCLLPQEWVQHFCQAPASQVDA 475

QY 473 LIRVNPDTGRVLFECCLKHSGYVTVAGTGHDPHDIVIPNGYFRFDSWNQFYTLAPMGNG 532

DB 476 LIRVNPDTGRVLFECCLKHSGYVTVAGTGHDPHDIVIPNGYFRFDSWNQFYTLAPMGNG 535

QY 533 AGRRR 537

DB 536 NGRRR 540

RESULT 8

AAB49710

ID AAB49710 standard; protein; 541 AA.

XX AC AAB49710;

XX





PR	22-JUN-1999;	99JP-00175928.	
XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.		
PA	(DENK-) DENKA SEIKEN KK.		
XX			
PI	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;		
XX			
XX	WPI; 2001-080848/09.		
DR	N-PSDB; AAF29141.		
XX			
XX	Kit for the detection and typing of small round-structured virus (SRSV)		
PT	strains for investigation of food poisoning outbreaks, contains		
PT	antibodies.		
XX			
PS	Claim 1; Page 40-42; 84pp; Japanese.		
XX			
CC	This invention relates to a kit for the detection and typing of small		
CC	round structured virus (SRSV) strains. The kit contains antibodies		
CC	directed against peptides represented in sequences AAB49700 - AAB49710,		
CC	which are each SRSV strain specific. Polynucleotide sequences AAF20141 -		
CC	AAF20151 represent cDNA encoding the strain specific proteins. The kit is		
CC	used for detecting and typing strains of SRSV in order to prevent the		
CC	spread of infection and to examine the epidemiology of outbreaks		
XX			
SQ	Sequence 545 AA;		
	Query Match 40.1%; Score 1157.5; DB 4; Length 545;		
	Best Local Similarity 45.6%; Pred. No. 2.3e-95;		
	Matches 253; Conservative 83; Mismatches 192; Indels 27; Gaps 14;		
Qy	1 MKMASNDANPS-DGST--ANLVPEVNN-EVMALPEVVGAAIAAPVAGQONVIDPWIRNMF 56		
Db	1 MMASKDAPTMDDTSGAGLVPEANTAEPIESVPAGAAATAAATAGQVNMIDPIMNNY 60		
Qy	57 VQAPGCEFTVSPNAPGEILWSAPGLDNLPLYSLHARMYNGYAGGFEVQVILAGNAFTA 116		
Db	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLAQMYNGVGNMKVLLAGNAFTA 120		
Qy	117 GKIIFAAVPPNPFTGELSPQVTMPHIIIVDVQLEPVLIPDPVRNNFHYNQSDSTI 176		
Db	121 GKIIISCIPGFAAQNISIAQATMFPHIADVRVLEPIEVLPEVDRNVLFH--NNDNAPTM 179		
Qy	177 KLIALMYTPLRA--NNAGDDVFTVSCRVLTPSPDFDFILVPTVESRTKPTVPILT 234		
Db	180 RLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFVLPNPVEQKTPESVNLPL 239		
Qy	235 EEMNSRPIPLEKLYTGPSSAFVVPQNGRCTTDDGVLGTTQLSAVNICTFRGDVTHIA 294		
Db	240 NTLNSRVPSLIIKSMVSRDHQVQFQNGRVTLDDGQLGTTPTTSASQLCKIRGSVFHAN 299		
Qy	295 GSHDYTMNLASQWNSNYDPTTEIPAPLGTDPFVGKIQGM--TOTTREDGSTRAHKATVS 352		
Db	300 GGNGY--NLTELDGSPYHAFES-PAPIGFPD-LGECDWHEASPTTQFTGDVVKQINVK 355		
Qy	353 TGSVHFTPKLSGVQYTTDTNNDFOGTQKTFPGVITODGNHQNHEPQQWVLPNYSRTG 412		
Db	356 QESA-FAPHLGTIQADGDSVSNVNTAKLGWSPVSDG--HRGDVDPWVIRYGSTLT 412		
Qy	413 HNVHLAPAVAPTFPGEQLLFFRSTMPGCGSYNNMNLCLLQEWVQHFQCEAAQAQSDVA 472		
Db	413 EAAQLAPPIYPGFGCEAIVFFMSDFIAHGTNGLSVPTCTIPQEFVTHFVNEQAFTGEAA 472		
Qy	473 LLRFVNPDGTVLFECKLHKSQYTV---AHTGPHDIVIPNGVFRFDSMVNQFYTLAP 528		
Db	473 LLHYLDPDTHRLGFBKLYPEGFMTCVNSGTPGQTL--PINGVFVSVNSRFYQLKP 530		
Qy	529 MNGA-----GRRRA 538		
Db	531 VGTAGPACRLGIRRS 545		
	RESULT 12		
	AAB49703		
ID	AAB49703 standard; protein; 544 AA.		
XX	AAB49703;		
AC			
XX	04-APR-2001 (first entry)		
DT			
XX	Small round structured virus protein SEQ ID 4.		
DE			
XX	Small round structured virus; SRSV; food poisoning.		
KW			
XX	Small round structured virus.		
OS			
XX	WO200079280-A1.		
BN			
XX	28-DEC-2000.		
PD			
XX	22-JUN-2000; 2000WO-JP004095.		
XX	22-JUN-1999; 99JP-00175928.		
PR	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.		
XX	(DENK-) DENKA SEIKEN KK.		
FA	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;		
PA			
XX	WPI; 2001-080848/09.		
PI	N-PSDB; AAF29144.		
DR			
XX	Kit for the detection and typing of small round-structured virus (SRSV)		
PT	strains for investigation of food poisoning outbreaks, contains		
PT	antibodies.		
XX			
PS	Claim 1; Page 47-49; 84pp; Japanese.		
XX			
CC	This invention relates to a kit for the detection and typing of small		
CC	round structured virus (SRSV) strains. The kit contains antibodies		
CC	directed against peptides represented in sequences AAB49700 - AAB49710,		
CC	which are each SRSV strain specific. Polynucleotide sequences AAF20141 -		
CC	AAF20151 represent cDNA encoding the strain specific proteins. The kit is		
CC	used for detecting and typing strains of SRSV in order to prevent the		
CC	spread of infection and to examine the epidemiology of outbreaks		
XX			
SQ	Sequence 544 AA;		
	Query Match 39.9%; Score 1151.5; DB 4; Length 544;		
	Best Local Similarity 44.6%; Pred. No. 7.9e-95;		
	Matches 254; Conservative 80; Mismatches 175; Indels 61; Gaps 16;		
Qy	1 MKMASNDANPS-DGST--ANLVPEVNN-EVMALPEVVGAAIAAPVAGQONVIDPWIRNMF 56		
Db	1 MMASKDAPTSADGATGAGQLVPEVNTADPIPDPVAGSSTALATAGOVNLDPWIINNF 60		
Qy	57 VQAPGCEFTVSPNAPGEILWSAPGLDNLPLYSLHARMYNGYAGGFEVQVILAGNAFTA 116		
Db	61 VQAPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMRVRVVLAGNAFTA 120		
Qy	117 GKIIFAAVPPNPFTGELSPQVTMPHIIIVDVQLEPVLIPDPVRNNFHYNQSDSTI 176		
Db	121 GKVIICCVPPGFSQRTSLIAQATLPHVIADVRLDPEVEPLEDRNVLYH--NNDTQPTM 179		
Qy	177 KLIALMYTPLRANNA--GDDVFTVSCRVLTPSPDFDFILVPTVESRTKPTVPILT 234		
Db	180 RLICMLYTLPTLRGASGTDSEVAVAGRVLTCPGDPDFLVPPTVEQKTRFTVPNPL 239		
Qy	235 EEMNSRPIPLEKLYTGPSSAFVVPQNGRCTTDDGVLGTTQLSAVNICTFRGDVTHIA 294		
Db	240 KYLSNSRIENPIEGMSLSPDQNTQVQFQNGRCTIDGQLGTPFVSVSQLCFKRGRIT--- 296		
Qy	295 GSHDYTMNLASQWNSNYDPTTEIPAPLGTDPFVGKIQGM--TOTTREDGSTRAHKA---TV 351		
Db	297 -SGQRVNLITELDGGSPF-MAPAAPAPAGFPD-LGSCDWHIEMSKIPNSSTQNPIVTNSV 353		
Qy	352 STGSVHFTPKL-----GSVQYTTDTNNDFOGTQKTFPGVITODGNH 395		

```

Db 354 KNSQOQVPHLSITLDENVSSGSDYIGTIQMTSPSD--SGGANTNF----- 399
Qy 396 QNEPOQWVLPNTSGRTGHNHVLAPAVAPTFPGEQLLFFRSTMPG--CSGYPMNNDCLLP 453
Db 400 -----WKIPDYGSSILAEASQAPAVYPGFNEVIVYFWASIPGPNQSGSNL--VPCLLP 452
Qy 454 QBWVHFCQBAAPASDVALLRFVNDPTGRVLFECKLHKSGYVTV-----AHTGPHDLVTP 509
Db 453 QBYITHFISEQAPIOGEAALLHYDPDTRNLGEFKLYPGGYLTCVPNSSSTGPOQL--P 510
Qy 510 PNGYRFESQWNOFYTLAPMGNGAGRRRAL 539
Db 511 LDGVVFASWSRFFYQLKPVGTAGPARGRL 540

RESULT 13
AAB49701
ID AAB49701 standard; protein; 530 AA.
XX
AC AAB49701;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 2.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 95JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
WPI; 2001-080848/09.
DR N-PSDB; AAF29142.
XX
PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
PS Claim 1; Page 42-45; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
SQ Sequence 530 AA;

Query Match 39.8%; Score 1148.5; DB 4; Length 530;
Best Local Similarity 44.8%; Pred. No. 1.4e-94;
Matches 247; Conservative 85; Mismatches 182; Indels 37; Gaps 14;

Qy 1 MKNASNDANPS--DGST--ANLYPEVN-NEVMALFEPVGAATAAPVAGQONVIDPWIRNF 56
Db 1 MWMASKDATSSVDGASGQLYPEVNASDPLAWDPVAGSSTAVATAGVNPDPWIINF 60
Qy 57 VOAPGGEFTVSPRNAPGEILWSAPLGPDLNPLYSLHARMYNGYAGGFVQVTLAGNAFTA 116
Db 61 VOAPQGEFTISPNNTPGGVLFDSLGLPHLNPFLHLSQWYNGWGVNMRVRIMLAGNAFTA 120

```

```

Qy 117 GKIIFAAVPPNPFTEGLSPSQVTMPPHIIVDVROLEPVLIPLDVRNNPFYHYNQNSDSTI 176
Db 121 GKIIIVSCIPPGGSHNLTAQATLPPHVIADVRLDPIEVPLEDRVNLHNNDRNQTM 180
Qy 177 KLIALMYTLPLRANNAGDDVFTVSCVLRTRPSDDPFIPLVPTVESRTPKFTVPILTVBE 236
Db 181 RLVCMLLYTLRLTGGGTGDSFVVAGRVMTCPSPDFNLFVLVPTVEQKTRPFTLPNLPUS 240
Qy 237 MSNSRFPILBLKLYTPSSAFVVOQNGRCTTGDVLLGTTQLSAVNICTFRGDVTHIAGS 296
Db 241 LNSRAPLPISGMGISPDNVQSVQFQNGRCTLGRLVGTTPVLSHVAKIRCT-----S 294
Qy 297 HDYTNLASQNSNYDPTTEETIPALGTDPFVCKIQGMLTQTTREDGSTRAHKATVSTGVS 356
Db 295 NGTVINLTDLGTTPHPPEG-PAPIGFPD-LGGCDWHINMT--QFGHSSQTQYDVTTPD 350
Qy 357 HFTPKLGSVOYTTDNDPQTGNTKFTPVGVIQ--DGNNHQNEPQ--QWVLPNYSGRGT 412
Db 351 FTVPHLSIQ-----ANGIGSNY-----IGVLSWVSPSPSHPSGSQVDLWKIPNYSST 400
Qy 413 HNVHLAPAVAPTFPGEQLLFFRSTMPGCGYPNMNLDCLLPQEWVQHFQCEAAPAQSDVA 472
Db 401 EATHLAPSVYPGFGEVLVFFWMSKIPGAY---SLPCLLPQEIYISHLASEQAPTVGEAA 457
Qy 473 LLRFVNPDTGRVLFECKLHKSGYVTV-----AHTGPHDLVIPPNGYFRFDSWYNQFYTLAP 528
Db 458 LLHYVDPDTGRTLGFKAYPDGFLTCVPGASSGPOQL--PINGVFVFWVSWRFPQLKP 515
Qy 529 MGNAGRRRAL 539
Db 516 VGTASSARGRL 526

RESULT 14
AAR57091
ID AAR57091 standard; protein; 530 AA.
XX
AC AAR57091;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 05-OCT-1994 (first entry)
XX
DE Small round virus SRSV/KY/89 capsid protein.
XX
KW pathogen; acute gastroenteritis; food poisoning; seafood contamination;
KW diagnostic assay; human calicivirus; small round virus; SRSV; KY89;
KW Norwalk virus; capsid protein.
XX
OS Small round structured virus.
XX
PN WO9405700-A2.
XX
PD 17-MAR-1994.
XX
PF 07-SEP-1993; 93WO-US008447.
XX
PR 07-SEP-1992; 92US-00941365.
XX
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI Matson DO, Bates MK, Jiang X, Graham DY;
XX
WPI; 1994-101125/12.
DR N-PSDB; AAQ56832.
XX
PT DNA from Norwalk and related viruses - used for preparing prods. for use
PT in diagnostic assays, detection and vaccines for Norwalk and related
PT viruses.
XX
PS Example 7; Fig 13a; 156pp; English.
XX

```





Db 524 RPYQLKPVGTASTARSL 541

Search completed: June 1, 2004, 13:45:52  
Job time : 48.3034 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.8419 Seconds  
(without alignments)  
2166.837 Million cell updates/sec

Title: US-09-926-799-5  
Perfect score: 2886  
Sequence: 1 MMASNDANPSDGSSTANLVP.....VNQFTYLPAGMGAGRRRAL 539

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pdp:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pdp:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pdp:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pdp:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pdp:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pdp:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1161.5	40.2	530	4	US-08-486-049-3
2	322	11.2	669	4	US-09-617-594A-2
3	303	10.5	623	4	US-09-590-020-4
4	300	10.4	623	4	US-09-590-020-2
5	299.5	10.4	626	4	US-09-590-020-7
6	298	10.3	622	4	US-09-590-020-6
7	297.5	10.3	668	4	US-09-617-594A-4
8	127	4.4	2318	3	US-09-091-219-24
9	127	4.4	2318	4	US-09-660-541-24
10	113	3.9	1042	4	US-09-252-991A-30444
11	111.5	3.9	776	3	US-08-961-083-86
12	111.5	3.9	776	4	US-09-536-784-86
13	109	3.8	607	4	US-08-556-422A-4
14	105.5	3.7	826	4	US-09-328-352-7515
15	104.5	3.6	1043	3	US-08-928-361B-30
16	104.5	3.6	1721	3	US-08-700-651-5
17	104.5	3.6	1721	3	US-08-928-361B-6
18	104.5	3.6	1721	4	US-09-588-995A-6
19	102.5	3.6	3892	4	US-09-328-352-5503
20	101	3.5	667	3	US-09-094-557-1
21	99.5	3.4	558	4	US-09-071-035-268
22	99.5	3.4	1638	4	US-09-071-035-258
23	99.5	3.4	1638	4	US-09-071-035-262
24	99.5	3.4	1638	4	US-09-071-035-266
25	99.5	3.4	1747	4	US-09-134-000C-5999
26	98.5	3.4	382	4	US-09-328-352-4587
27	98.5	3.4	915	1	US-08-328-322-5

28	98.5	3.4	1013	3	US-09-415-522-8	Sequence 8, Appli
29	98.5	3.4	2736	4	US-09-252-991A-30227	Sequence 30227, A
30	98.5	3.4	2972	4	US-09-579-181-2	Sequence 2, Appli
31	98.5	3.4	3118	4	US-09-579-181-1	Sequence 1, Appli
32	98	3.4	627	4	US-09-328-352-7547	Sequence 7547, Ap
33	97.5	3.4	1042	3	US-08-928-361B-11	Sequence 11, Appl
34	97.5	3.4	1042	4	US-09-588-995A-11	Sequence 11, Appl
35	97.5	3.4	1837	3	US-08-928-361B-5	Sequence 5, Appli
36	97.5	3.4	1837	4	US-09-588-995A-5	Sequence 5, Appli
37	96	3.3	1075	5	PCT-US94-07297-41	Sequence 41, Appl
38	95.5	3.3	1091	6	5516630-2	Patent No. 5516630
39	95.5	3.3	1810	4	US-08-793-273C-4	Sequence 4, Appli
40	95.5	3.3	1810	5	PCT-US95-11684-4	Sequence 4, Appli
41	95.5	3.3	2227	3	US-08-475-886-2	Sequence 2, Appli
42	95.5	3.3	2227	3	US-08-475-886-4	Sequence 4, Appli
43	95.5	3.3	2227	3	US-08-475-886-6	Sequence 6, Appli
44	95.5	3.3	2227	3	US-08-397-232-2	Sequence 2, Appli
45	95.5	3.3	2227	3	US-08-397-232-4	Sequence 4, Appli

#### ALIGNMENTS

RESULT 1  
US-08-486-049-3  
; Sequence 3, Application US/08486049  
; Patent No. 6572862  
; GENERAL INFORMATION:  
; APPLICANT: Bates, Mary K  
; APPLICANT: Jiang, Xi  
; APPLICANT: Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE:  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-486-049-3

Query Match 40.2%; Score 1161.5; DB 4; Length 530;  
Best Local Similarity 45.4%; Pred. NO. 5.6e-104;  
Matches 250; Conservative 83; Mismatches 181; Indels 37; Gaps 14;  
QY 1 MMASNDANPS-DGST--ANLVPEVN-NEVMALPEVVGAAIAPVAGQGVDPHNNF 56  
DB 1 MMASKDTSVDGASGAGQLVPEYNASDDPLAMDVPAGSVATAVAGQVNPDPINN 60

QY 57 VOAPGGETVSPRNAPGAILMSAPLGPDLNPNYLSHLARMYNGYAGGEVQVVLGNAGFTA 116  
DB 61 VOAPGGETVSPRNAPGAILMSAPLGPDLNPNYLSHLARMYNGYAGGEVQVVLGNAGFTA 120  
QY 117 GKIIIFAAVPPNPPTGEGSPQVTPMPPHIIIVDRQLPVLIIPLDPVRNNFYHYNQSDSTI 176  
DB 121 GKIIIVSCITPPGGSHNLIIAQATLFPHVIAVRLTDPIEVLDPVRLFNHNDNRNQTM 180  
QY 177 KLIAMLYPLRANAGDDVFTVSCVLTTRPSPDFIFLVPPTVESRTKPTVPLTVVEE 236  
DB 181 RLVCMLYTLRTGGGTGDSFVAVGRVMTCPSPDFNLFVPTVEQKTRPTLLENPLSS 240  
QY 237 MNSRPPIPLEKLYTGPSSAFVVPQNGRCTTGDVLLGTQLSAVNICTRFGDVTHIAGS 296  
DB 241 LNSRAPLPISSMGISPDNVQSVQFQNGRCTLDGLRGVTPVLSHVAKIRGT-----S 294  
QY 297 HDVTNMLASQNSNYDPTTEIPAPLGTDFVKGIOQMLTQTTREDSGSTRAHKATVSTGSV 356  
DB 295 NGTVINLTDLGTPPHFEG-PAPIGFPD-LGGCDWHINMT--QFGHSSQTQYDVTTPD 350  
QY 357 HPTPKLGSVQYTTDNDTQNTKFTVGVIIQ--DGNHONPEQ--QWVLPNYSGRGT 412  
DB 351 TFVPHLGSIQ-----ANGISGNY-----VGLSWISPSHSGSQVDLWKIPNYGSSIT 400  
QY 413 HNVHLAPAVAPTPFGGQQLFFRSTMPGCSGYFNMNLDCLLPQEWVQHFCEAAPASQDVA 472  
DB 401 EATHLAPSVYPPGFEVLFNFKMFGPGAY---NLPCLLPQEIYSHLASEQAPTVEAA 457  
QY 473 LLRFVNPDTGRVLFECMLKHSQVTV-----AHTGPHDLVIPPNGYFEREDSWNOFTLAP 528  
DB 458 LLHVYDPDTGRNLGEKAYPDGLTFCVPGASSGPOOL--PINGVVFVSWVSFRFQLKP 515  
QY 529 MNGGAGRRAL 539  
DB 516 VGTASSARGRL 526

RESULT 2  
US-09-617-594A-2  
; Sequence 2, Application US/09617594A  
; Patent No. 6541458  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.1  
; CURRENT APPLICATION NUMBER: US/09/617,594A  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 669  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-617-594A-2

Query Match 11.2%; Score 322; DB 4; Length 669;  
Best Local Similarity 25.4%; Pred. No. 3,1e-22;  
Matches 125; Conservative 69; Mismatches 158; Indels 140; Gaps 22;

QY 11 SDGSTANLVPEVNVEV--MALEPVVGAATAAPVAGQNVDPWIRNPNVQAPGGEFTV-- 66  
DB 125 AUGDSITTEQGTGLVGGVIAEPSAQMTAATAATKSKVDSEW--ESF-----FSFHT 175  
QY 67 ----SPRNAPGAILMSAPLGPDLNPNYLSHLARMYNGYAGGEVQVVLGNAGFTAAGKIIIFA 122  
DB 176 SVNWSTSETQKILFKQSLGPLNPNYLHLSKLYVANWGSVDVRFSGSGVFGGKLAAI 235

QY 123 AVPPNPPTGEGSPQVTPMPPHIIIVDRQLPVLIIPLDPVRNNFYHYNQSDSTIKLI 179  
DB 236 VVPP-----GVDFVQSTSMQLQPHVLFDAQVEPVIFSIPLDLRSTLYHLMSDITDT-SLV 289  
QY 180 AMLYTPLRANAGDDVFTVSC--RVLTTRPSPDFIFLVPPTVESRTKPTVPLTVVEE 222  
DB 290 IMVNDL-INPYANDSNSSGCVTVVETKPGDFKPHLLKPPGSMMLTHGSIPLSDLIPKSSS 348  
QY 223 -----RTKPTVPLTVVEE----- 243  
DB 349 LWIGNRYMSDITDFVIRPFVFOANRHDFNQETAGWSTFRFRPTI--TIESNGS---- 402  
QY 244 IPLEKLYTGPSSAFVQV--PQNGRCTTGDVLLGTQLSAVNICTRFGDVTHIAGSDHYTM 301  
DB 403 -----KLGTVATDIYIVPGIPDGPWPTTIGEEEL-----TPAGDYSITNGSGN--- 444  
QY 302 NLASQNSNYDPTTEIPAPLGTDFVKG-IOQMLTQ-----TTREDSGSTR 346  
DB 445 DIATAN--AYDSADVI---TNTTFRGMVTCALQARAWGDKKISSTAFITTAKEGNTLK 499  
QY 347 HKATVTCGSHVHTPKLGSVQYTTDNDTQNTKFTVGVIIQ--DGNHONPEQ--QWVLPNYSGRGT 412  
DB 500 PSNTIDMTKI-----AVYQDTHVGRDVQTSDDTLAILGYTGIGEAIGSNRDSVVRIS 552  
QY 403 VLPNYSGRGTGHN 414  
DB 553 MLPETGARGGNH 564

RESULT 3  
US-09-590-020-4  
; Sequence 4, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Aivars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; PRIOR FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-4

Query Match 10.5%; Score 303; DB 4; Length 623;  
Best Local Similarity 26.1%; Pred. No. 1.9e-20;  
Matches 128; Conservative 55; Mismatches 167; Indels 140; Gaps 23;

QY 12 DGSTANLVPEVNVEV--MALEPVVGAATAAPVAGQNVDPWIRNPNVQAPGGEFTV-- 66  
DB 82 DGSIT--TPEQGTMTVGGVIAEPSAQMTAATAATKSKVDSEW-----EAFSFTSVNM 133  
QY 67 SPENAPGAILMSAPLGPDLNPNYLHLSKLYVANWGSVDVRFSGSGVFGGKLAIVVPP 126  
DB 134 STSETOKILFKQSLGPLNPNYLHLSKLYVANWGSVDVRFSGSGVFGGKLAIVVPP 193  
QY 127 NFPTGEGSPQVTPMPPHIIIVDRQLPVLIIPLDPVRNNFYHYNQSDSTIKLI 183  
DB 194 -----GIEPVQSTSMQLQPHVLFDAQVEPVIFSIPLDLRSTLYHLMSDITDT-SLVIMVY 247  
QY 184 TPLRANAGDDVFTVSC--RVLTTRPSPDFIFLVPPTVESRTKPTVPLTVVEE 222  
DB 248 NDL-INPYANDTNSSGCVTVVETKPGDFKPHLLKPPGSMMLTHGSIPLSDLIPKSSS 306

```
QY 223 -----RTKPTVPILTVEMSNSRPPIPLE 247
Db 307 NRHWSIDTDFIIRPFVQANRHFDFNQETAGMSTRFRFRDITI---TVSE-SN-----MS 356
QY 248 KLYTGPSSAFVQ--PONGRCCTDGLVLLGTQLSAVNICTFRG-DVTHIAGSHDYTMNLA 304
Db 357 KLGIGVATDIYVPGIPDGWPDITIEQL--TPAGIYSITASNGTDITTAAG-----405
QY 305 SONWNSYDTEBIPAPLPGPDF-----VGKIQ-----GMLTQTTREDGSTRAHK 348
Db 406 -----YDAAETI---VNTTFKSMYICGSLQRAWGDKKISNTAFITTAVRKGSIEPSN 456
QY 349 ATVSGSVHFTPKLGSVOYTTDNDPFGQNT-----KFTPGVLODGNHQNHPQQWVL 404
Db 457 TIDMTKL-----VYODAHVGEVOTSDITLALLGYTGIGEBAGSDRDKVVRISVL 508
QY 405 PNYSGRTGHN 414
Db 509 PETGARGGNH 518

RESULT 4
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-2

Query Match 10.4%; Score 300; DB 4; Length 623;
Best Local Similarity 26.9%; Pred. No. 3.7e-20;
Matches 97; Conservative 45; Mismatches 113; Indels 106; Gaps 14;

QY 12 DGSSTANLVPEVNNV--MALEPVVGAATAAPVAGQONVIDPWIRNFFVOAPGGEFTV---66
Db 82 DGSIT--TPEQGTWVGVIAPESQAQMSAADMATKSDSEW-----EAFPSFHTSVNW 133
QY 67 SPRNAPGELLMSAPLGPDLNPLYSLARMYAGVGFVQVILAGNAPTAGKIIFAAVPP 126
Db 134 STSETQKILFKQSLGPLNPLYSLAKLYVAMSGSIEVRFSIGSGVFGGKLAIVVPP 193
QY 127 NPPTGLSPSQVT---MFFHIIVDVRLPEVLPILPDVNRNFFHYNQSNDSIKLIAMLY 183
Db 194 -----GIEPVQSTMLQYPHVLFDAQVEPVIFALPDLRSNLYHLMSDITTT-SLVIWY 247
QY 184 TPLRANNAGDDVFTVSC--RVLTRSPDPDFILVPPVSTKPTFTVPILTVEMSNSR 241
Db 248 NDL-INPVANDTNSSGCIIVTETKGPDPKFHLLKPP--GSMITHGSPVSDIIPKSS--302
QY 242 PPIPLEKLYTGPSSAFVQVQ---QNGR-----CTTDGVLGTTQLSAVNICTFRGDVTHI 293
Db 303 --LWIGNRHSIDITDFIIRPFVQANRHFDFNQETAGMSTRFRFRDITI---TVSE-SN 356
QY 294 AGSHDYTMNLSAQNSNYD--PTEBIPAPLGTDPFVGKIQGMLTQTTREDGSTR--HKAT 350
Db 361 GVATDIYVPGIPDGWPDITIEQLPA-----GDYA-----ITNGTGNDITATGYDTAD 410
QY 351 VSTGSVHF-----TPKLGSVQVYTT-----DTNNDPQTQNTKFTVPVGIQ 390
Db 411 IIKNTNFRGMWYICGSLQRAWGDKKISNTAFITATLGDNNKINPCNTIDQSKIIVVQ 470
QY 391 D---GNHQNHPQQWVLNYSQ-----RTCHNVHLAPAVAPTP-----GEQLLFRSTM 437
Db 471 DAHVKKQAQTSDDTLALLGYTGIGSQAIGSDRDRVVRIS-TLPETGARGGNHPIFYKSI 529
QY 438 PGCSYPPNNMLDCLLPQEWQHFQCEAA-----PAQSDVALLRFPVNPDTGRVLFECKLH 491
Db 530 K--LGYVTRSDVFNQ--ILHTRQSLNHYLLPDPDS-FAYVRII--DSNGSWFDIGD 582
QY 492 KSGYVTVVAHTGPHDLVIP 509
Db 583 SDGFSFVGSGFKLEFP 600

RESULT 6
```

```
QY 223 -----RTKPTVPILTVEMSNSRPPIPLE 247
Db 307 NRHWSIDTDFIIRPFVQANRHFDFNQETAGMSTRFRFRDITI---TVSE-SN-----MS 356
QY 248 KLYTGPSSAFVQ--PONGRCCTDGLVLLGTQLSAVNICTFRG-DVTHIAGSHDYTMNLA 304
Db 357 KLGIGVATDIYVPGIPDGWPDITIEQL--TPAGIYSITASNGTDITTAAG-----405
QY 305 SONWNSYDTEBIPAPLPGPDF-----VGKIQ-----GMLTQTTREDGSTRAHK 348
Db 406 -----YDAAETI---VNTTFKSMYICGSLQRAWGDKKISNTAFITTAVRKGSIEPSN 456
QY 349 ATVSGSVHFTPKLGSVOYTTDNDPFGQNT-----KFTPGVLODGNHQNHPQQWVL 404
Db 457 TIDMTKL-----VYODAHVGEVOTSDITLALLGYTGIGEBAGSDRDKVVRISVL 508
QY 405 PNYSGRTGHN 414
Db 509 PETGARGGNH 518

RESULT 4
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-2

Query Match 10.4%; Score 300; DB 4; Length 623;
Best Local Similarity 26.9%; Pred. No. 3.7e-20;
Matches 97; Conservative 45; Mismatches 113; Indels 106; Gaps 14;

QY 12 DGSSTANLVPEVNNV--MALEPVVGAATAAPVAGQONVIDPWIRNFFVOAPGGEFTV---66
Db 82 DGSIT--TPEQGTWVGVIAPESQAQMSAADMATKSDSEW-----EAFPSFHTSVNW 133
QY 67 SPRNAPGELLMSAPLGPDLNPLYSLARMYAGVGFVQVILAGNAPTAGKIIFAAVPP 126
Db 134 STSETQKILFKQSLGPLNPLYSLAKLYVAMSGSIEVRFSIGSGVFGGKLAIVVPP 193
QY 127 NPPTGLSPSQVT---MFFHIIVDVRLPEVLPILPDVNRNFFHYNQSNDSIKLIAMLY 183
Db 194 -----GIEPVQSTMLQYPHVLFDAQVEPVIFALPDLRSNLYHLMSDITTT-SLVIWY 247
QY 184 TPLRANNAGDDVFTVSC--RVLTRSPDPDFILVPPVSTKPTFTVPILTVEMSNSR 241
Db 248 NDL-INPVANDTNSSGCIIVTETKGPDPKFHLLKPP--GSMITHGSPVSDIIPKSS--302
QY 242 PPIPLEKLYTGPSSAFVQVQ---QNGR-----CTTDGVLGTTQLSAVNICTFRGDVTHI 293
Db 303 --LWIGNRHSIDITDFIIRPFVQANRHFDFNQETAGMSTRFRFRDITI---TVSE-SN 356
QY 294 AGSHDYTMNLSAQNSNYD--PTEBIPAPLGTDPFVGKIQGMLTQTTREDGSTR--HKAT 350
Db 361 GVATDIYVPGIPDGWPDITIEQLPA-----GDYA-----ITNGTGNDITATGYDTAD 410
QY 351 VSTGSVHF-----TPKLGSVQVYTT-----DTNNDPQTQNTKFTVPVGIQ 390
Db 411 IIKNTNFRGMWYICGSLQRAWGDKKISNTAFITATLGDNNKINPCNTIDQSKIIVVQ 470
QY 391 D---GNHQNHPQQWVLNYSQ-----RTCHNVHLAPAVAPTP-----GEQLLFRSTM 437
Db 471 DAHVKKQAQTSDDTLALLGYTGIGSQAIGSDRDRVVRIS-TLPETGARGGNHPIFYKSI 529
QY 438 PGCSYPPNNMLDCLLPQEWQHFQCEAA-----PAQSDVALLRFPVNPDTGRVLFECKLH 491
Db 530 K--LGYVTRSDVFNQ--ILHTRQSLNHYLLPDPDS-FAYVRII--DSNGSWFDIGD 582
QY 492 KSGYVTVVAHTGPHDLVIP 509
Db 583 SDGFSFVGSGFKLEFP 600

RESULT 6
```

us-09-926-799-5.ra1

Wed Jun 2 09:13:29 2004

```

; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-617-594A-4

Query Match      10.3%; Score 297.5; DB 4; Length 668;
Best Local Similarity 26.8%; Pred. No. 7.3e-20;
Matches 110; Conservative 52; Mismatches 147; Indels 101; Gaps 20;

QY 12 DGSTANLVPEVNNEVMALEPVVGAIAAPVAGQONVIDPWRNNFVQAPGGEF-----64
DB 127 DGSVT--TPEQGT-----LVGGVIAEPNA-QMSAVADVATGKSDSEWAEFFSFHTSV 176
QY 65 TVSPRNAPGEIILWSAPLGDPLNLYSLHARMYNGVAGFEVQVILAGNAFTAGKIIFAAV 124
DB 177 NWSTSETQKILFKQSLGSLPLNLYSLHAKLVANGSGSIEVRFSGSGVFGKLAIVV 236
QY 125 PPNFTEGLSPSOVT---MFPHIIVDVRLQLEVLPLPDVRNNFHYNQSDSTIKLIAM 181
DB 237 PP-----GIDPVQSTSLMLQYPHVLFDAQVPEVFIPTDLRNSLYHLMSDITDTT-SLVIM 290
QY 182 LYTPLRANNAGDDVFTVSC--RVLTSPSPDFEFLVPPPTVESRTKPTFVPIITVEEMSN 239
DB 291 IYNDL-INPYANDNSNGCIVTVETKPGDPDFKHLKPPG-----SMLTHGSIFS 339
QY 240 SRPIPLEKLYTGPSSAFVVPQNGRCTTDDGVLGTTQLSAVNICTRGDVTHIAGSHDY 299
DB 340 DLIP-----KSSSLWIGNRHWSDIITDFV-----IKPFVFQ-----ANRH-F 374
QY 300 TMNLASQWNSNYDPTEIPAPLGTDPFVGKIQGMLTQTTREDGSTRAHKATVSTGSV---356
DB 375 DFNQETAGWS-----TPRF-----RPITITVSEKGSKL-GIGVATDSIVPG 415
QY 357 -----HFTPKLGSVQVTTDNDFTQNTKFTFVGVQIDGNN 394
DB 416 IPDGMPTTPEIKLTP-AGDYAITNGGNDITTAAD--YDGASIIKNTN 462

RESULT 8
US-09-091-219-24
; Sequence 24, Application US/09091219
; Patent No. 6171592
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABE, Brendan S.
; APPLICANT: PENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/091,219
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: PCT/AU96/00815
; EARLIER FILING DATE: 1996-12-18
; EARLIER APPLICATION NUMBER: AU PN7201
; EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-09-091-219-24

Query Match      4.4%; Score 127; DB 3; Length 2318;
Best Local Similarity 22.3%; Pred. No. 0.021;
Matches 73; Conservative 39; Mismatches 134; Indels 82; Gaps 14;

QY 92 LARMYNGYAGGFEVQVILAGNAFTAGKIIFAAVPPNFPTEGLSPSOVTMFPHIIVDVRLQ 151
DB 380 LTDSYAMRNGWDVEVTAVGNQNGCLLVAMVPYLSIQKRELYQLTLFPHQFINPRTN 439
QY 152 EPVLIPLPDVRNNFY-HYNQSDSTIKLIAMLYTPLRANNAGDDVFTVSCRVLTRPSPDF 210

```

Db 440 MTAHITVPFVGNRYDQYKVKHPWT--LVVMVAPLTVNTEG-----APQI 483  
QY 211 DFILVPPVESRTPFTVPILTVEMSNRRPPIPLEKLY-----TGPSSAFV----- 259  
Db 484 KYVANIAPT-----NVHAGEFSPKEGIFPVACSDGYGLVTTDPKTDADPVYGVKFN 535  
QY 260 QPON--GRCITDGLVLTQTSANVNICFRGDVTHIAGSHD-----YTNLASQNS 309  
Db 536 PPRNQLPGRFTN---LLDVAE-ACPTFLRFEFGVPYVTTKTDSDRVLQAQFDMSLAOKMS 591  
QY 310 NYDPTPEEIPAPLGTDFVKGICQMLTQTTREDGSTRAHKATVSTG-----SVHFTPK 361  
Db 592 N-----TFLAGLAQYTYQYSGTINLH--FMFTGPTDAKARYMWAYAPP 632  
QY 362 LGSVQYTTDT-----NNDFQTQNTKFT 384  
Db 633 GMEPPKTPAAAHCIAHAEWDTGLNSKFT 660

RESULT 9  
US-09-660-541-24  
; Sequence 24, Application US/09660541  
; Patent No. 6531136  
; GENERAL INFORMATION:  
; APPLICANT: STUDDERT, Michael J.  
; APPLICANT: CRABB, Brendan S.  
; APPLICANT: FENG, Li  
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS  
; FILE REFERENCE: 040268/0151  
; CURRENT APPLICATION NUMBER: US/09/660,541  
; CURRENT FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/091,219  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201  
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 2318  
; TYPE: PRT  
; ORGANISM: Foot-and-mouth disease virus  
US-09-660-541-24

Query Match 4.4%; Score 127; DB 4; Length 2318;  
Best Local Similarity 22.3%; Pred. No. 0.021;  
Matches 73; Conservative 39; Mismatches 134; Indels 82; Gaps 14;

QY 92 LARMYNGAGFEVQVILAGNAFTAGKIIFAAVPPNFPFTEGLSPSQVTMFPHIIVDVRLQ 151  
Db 380 LTDSAYMRNGWDVETAVGNQFNGCGLLVAMPPELYSIQKRELYQLTLFPHQFINPRTN 439  
QY 152 EPVLIPLDVRNNFY-HYNQSNDSITKILAMLYPLRANNAGDDVFTVSCRVLTRPSPDF 210  
Db 440 MTAHITVPFVGNRYDQYKVKHPWT--LVVMVAPLTVNTEG-----APQI 483  
QY 211 DFILVPPVESRTPFTVPILTVEMSNRRPPIPLEKLY-----TGPSSAFV----- 259  
Db 484 KYVANIAPT-----NVHAGEFSPKEGIFPVACSDGYGLVTTDPKTDADPVYGVKFN 535  
QY 260 QPON--GRCITDGLVLTQTSANVNICFRGDVTHIAGSHD-----YTNLASQNS 309  
Db 536 PPRNQLPGRFTN---LLDVAE-ACPTFLRFEFGVPYVTTKTDSDRVLQAQFDMSLAOKMS 591  
QY 310 NYDPTPEEIPAPLGTDFVKGICQMLTQTTREDGSTRAHKATVSTG-----SVHFTPK 361  
Db 592 N-----TFLAGLAQYTYQYSGTINLH--FMFTGPTDAKARYMWAYAPP 632  
QY 362 LGSVQYTTDT-----NNDFQTQNTKFT 384  
Db 633 GMEPPKTPAAAHCIAHAEWDTGLNSKFT 660

RESULT 10

US-09-252-991A-30444  
; Sequence 30444, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE REFERENCE: 107186.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 30444  
; LENGTH: 1042  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-30444

Query Match 3.9%; Score 113; DB 4; Length 1042;  
Best Local Similarity 22.8%; Pred. No. 0.13;  
Matches 126; Conservative 53; Mismatches 234; Indels 140; Gaps 28;

QY 11 SDGSTANLVPEVNEVMALEPVVGAIAAPVAGQONVI-----DPWIRNFFVAPGG 62  
Db 21 ADGATESTV-----VLAGPLALPAASSATALITVPLASGVGVK---VOLPLP 65  
QY 63 EFTVSPRNPAGEILLMSAPLGDLPNPLS-HLARMYNGVAGGFEVQVILAGN-AFTAGKII 120  
Db 66 SAVVSPIGLPLSPVRITLAPTSAVEVITLPLLLALITGVAGATVSTVVLGVLPLPAGS-- 123  
QY 121 FAAVPPNPFTEGLSPSQVTMFPHIIVDVRLQEPVLIPLDVRNNFYHYNQSNDSITKILIA 180  
Db 124 WATALITVPLASGVGVKVLQPL---LPSARVWPVIGLPLSVR-----ITRTGSAVPL-- 173  
QY 181 MLYTPLRANNAGDDVFTVSCRVLTRPSPDFFIPLVPPTVESRTPKFTVPILTVBEMSNS 240  
Db 174 -SLTPLLGSITGVAGATASTVLTGP-----LVLPAASWATALITVPLVSGVGVK 224  
QY 241 RPIPLEKLYTGPSSAFVQVQNGRCITTDGVLGTLTQTSANVNIC--FRGDVTHIAGSHD 298  
Db 225 QLPLP-----SAVACRIGLPPSPVRIT-----VAPASAVPLITVPLLGSSTGVAGATE 272  
QY 239 YTMILA-----SQWNSNDPTEEIPAPLGTDFVKGICQMLTQTR-----EDGSTRAH 347  
Db 273 STVVLGAPLVLPAASWASALITE--PLASGVGV--KLQLPLPSAVTWPIGLPPSPVRIT 328  
QY 348 KATVSTGSHVFTPKLGSVQYTTDNNDFQTQNTKFTPVGVIQDGNHNEPQQMWLPNY 407  
Db 329 VAPASAVPAITLPLLG-----LTTG-----V 349  
QY 408 SGRTGHNVLH-APAVAPTFGEQLLFFRSTWPGSGYNNMLDCLLPQ--EWWQHFQCEA 464  
Db 350 AGATASTVVLVGPVLPLPAGSWATALI---TVPLASGVGVKQLQLPLPSAVAMPVIGLPPS 406  
QY 465 -----APAQSDVALLRFVNPDTGRVLFECCLKHSGYVTVVAHTGPHDLVIPPNGYFRFDSW 519  
Db 407 VRITVAPASAVPAI---TLPLLGSITGAAGATES---TVVLAPG--LVLPA-----ASW 452  
QY 520 VNQFYTLAPMGNG 532  
Db 453 ATAL-TTVPLDSDG 464

RESULT 11  
US-08-961-083-86  
; Sequence 86, Application US/08961083  
; Patent No. 6159469  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines





Db 256 DGK-----VPAVTA--SVSNGLATVVPVREGEPRVRIAKAENGDIILGEYRLHFTKOKSL 309  
Qy 273 LGTTOLSAVINICTER--GDV-----THIAGSHDYTNILASQNSVNDPTEIPAPLG 322  
Db 310 LSHKFAVAAKQARLLQVQALELPTKVPVYFGKGDTKDLTVE-----EEVPAENL 363  
Qy 323 PTDVFGKIQGMLTQTTREDGSTRAHKATVSTGSHVHTPKLGSV-----QYTTDTNNDPQT 377  
Db 364 TK-----AGQFTVRGRVLGSLVAEIT-----VRVTDKLGTELSDNPNYDENSNAQAFAS 412  
Qy 378 CON-----TKFPTVGVIOQGNHNEPQOWVLNPNYSGRTGHNHVLAPAVAPFPQGLLFF 433  
Db 413 ATNDIDKNSHRVDYLDGDHSEN--RRWT--NWSPTSSNEVSAGV-----IF 458  
Qy 434 RSTMPGCSGYPNMNLDCLLPQWVQHFCQEAAPQSDVALLRFVNPDTGRVLFECCLHKS 493  
Db 459 REN-----GKIVERVTQGVQKQVFPADSGTDAFSLVLEYVGP-----FEVPTYYS 505  
Qy 494 GYVTVVAHTGPHDLVTPPN 511  
Db 506 NY--QAYDADHPFNPN 521

## RESULT 13

US-08-556-422A-4  
; Sequence 4, Application US/08556422A  
; Patent No. 6576754

## GENERAL INFORMATION:

; APPLICANT: HALL, Kathryn T.  
; APPLICANT: FREEMAN, Gordon J.  
; APPLICANT: SCHULTZE, Joachim L.  
; APPLICANT: BOUSSIOTIS, Vassiliki  
; APPLICANT: NADLER, Lee M.  
; TITLE OF INVENTION: NUCLEIC ACIDS ENCODING CD100 MOLECULES  
; FILE REFERENCE: DFN-005CPA2  
; CURRENT APPLICATION NUMBER: US/08/556,422A  
; CURRENT FILING DATE: 1995-11-09  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 607  
; TYPE: PRT  
; ORGANISM: Mus musculus

US-08-556-422A-4

Query Match 3.8%; Score 109; DB 4; Length 607;  
Best Local Similarity 18.9%; Pred. No. 0.13;  
Matches 106; Conservative 59; Mismatches 165; Indels 230; Gaps 29;  
Qy 52 IRNNEVQAPGGEFTVSPRNAPGEILWSAPLG-----ELLSADADRKKQCSFKGDKPKDCQNYIKILLPINS 98  
Db 39 LNSNLSFLPGGEYQ-----LLMSADADRKKQCSFKGDKPKDCQNYIKILLPINS 90  
Qy 99 YAGGEVQVILAGNA-----FT-----AGKIF-----AAPPNF----- 128  
Db 91 H-----LLTCGTAAPLCAVIAHSFTLAQDEAGNVILEDKGHCPCFPDFNFKSTALV 143  
Qy 129 -----PTEGLSPQVTFPHIIIVDVQLEPVLIPDV 161  
Db 144 VDGELYTGVTSFQGNDFPAISRSQRPTKTESSLNWLQDPAFASATSPESLGSPIGDD 203  
Qy 162 RNNFVHYNQSN-----DSTIKLIAMLYTPLRANNAGDDV-----FTVSCRVLTRSP 208  
Db 204 DKYIFFSETGCEPPEPNTI--VSRVARVCKDEGGERVLOQRWTSFLKQALLCSRDD 261  
Qy 209 PDFDFILVPTVESRTKPTFTVILFVEEMSNRFPPIPLEKLYT--GPSSAFVVPQNGRGT 267  
Db 262 GF-----PENV-----LQDVFTLNP-----PDWRKT 284  
Qy 268 TD-GVLL-----GTTQLSAVNICTF--RGDVTHIAGSHDYTNILASQNSVNDPTEIPAP 320  
Db 285 LSIQVTSQWHRGTTEGSA--ICVETMNDVQKAFDGLYKKVNRETQW--YTETHQVPTP 340

Qy 321 -----LGTDPDFVGKIQGMLTQTTREDGSTRAH-----KATVSTGVS 356  
Db 341 RPGACITNSARERKINSSQLPD---RVNLFKDHLMDGOVRSRLLLQLQPRARYQVAV 397  
Qy 357 HFTPKLGSVQVYTTDNDPQTGQNTKFTPVGVIOQGNH-----QN 397  
Db 398 HRVFGHLST-----YDVLFGTG---DGRLHKAVTLSSRVHIIIELOIFPQG 441  
Qy 398 EP-QQWVLNPNYSGRTGHNHVLAPAVAPFPQGLLFFRSTMPGCSGYPNMNLDCLLPQB- 455  
Db 442 QPVQNLLDLSHGGLLYASSHSGVQVP-----VANCSLYPTCG--DCLLARDP 487  
Qy 456 ---VQHFCQBAAPQSDVA 472  
Db 488 YCAWTSACRLASLYQPDLA 507

## RESULT 14

US-09-328-352-7515  
; Sequence 7515, Application US/09328352  
; Patent No. 6562958

## GENERAL INFORMATION:

; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; FILE REFERENCE: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7515  
; LENGTH: 826  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii

US-09-328-352-7515

Query Match 3.7%; Score 105.5; DB 4; Length 826;  
Best Local Similarity 22.7%; Pred. No. 0.47;  
Matches 116; Conservative 60; Mismatches 180; Indels 155; Gaps 32;  
Qy 58 QAPGGEFTVSPRNAPGEILWSAPLGDLNPLYSLHARMYGVAGGFEVQVILAGNAFTAG 117  
Db 359 KAPREFTVSAPTIACTIIVGSRADNV-----AADMPPGVIRGYDV-----ITG 403  
Qy 118 KIIFAAVP-----PNF-----PTEGLSPSQVTFPHIIIVDVQLEPVLIPL-----PDV-----R 162  
Db 404 KLRWAFDPNPDNPNVYLKGEIYKRSSANSWAAMSVD--PQNTVFLPMGSSSVDIWGNR 462  
Qy 163 NNFVH--YNQS---NDSTIKLIAMLYTPLRANNAGDDVFTVSCRVLTRSPDPDFIF--L 215  
Db 463 NPLDHKYNSTVLALDATTGKEKVVYQTVH-----NDLW-----DPDLPMQPSL 505  
Qy 216 VP-PTVESRTKPTFTV-----PILTVEEMSNRFPPIPLEKL--YTGPSA 256  
Db 506 VDFPLKGGTTKPAVVGITKSGQFFVLDRVTKPLTKVIEQVPKAAIIPGEQVSLTPQRS- 564  
Qy 257 FVWQPNQGRCT--TDGVLILGTQ---LSAVNICTPRGDVTHIAGSHDYTM-----NLASQ 307  
Db 565 -VENPQIGNQTLTESDMKGATFPDQLMCRINFKSNRYEGLTAPCTDVSLSPPGSLGGNN 623  
Qy 308 WSN--YDTEIIPAPLGPDPVKGKIQ--GM-----LQTTREDGSTR-----HKATVSTGSHV 358  
Db 624 WGSTAFDPDTHRYM-----FVNDMLGLMWQLIKQTPEDIKIQASGCKVKNTGMGAV-- 674  
Qy 359 TPKLGS-----VQYTTDNDPQTGQNTKFTPVGVIOQGNHNEP 399  
Db 675 -PMKGTPIYKVNKRPMFMSALSIPOCKPPFGTMTAIDMKTRQVAVQVPLGTVEB----- 725  
Qy 400 QQWVLEPNYSGRTGHNHVL--APAVAPTPGEQ-----LLFFRSTMPGCSGYPNMNLDCLLP 453  
Db 726 -----TGPLGLKMLKAPIGHPTTGGPMATOGGLVFFAAT-----QDYILRAFDSTG 773  
Qy 454 QE-VQHFCQBAAPQSDVALLRFVNPDTGR 483

Qy 242 FPIPLEKLYTGSSAFVVPQNGRCCTTGDVLLGTTQLSAVNICTFRGDVTHIAGSHDYTM 301  
Db 303 ---PTTGLPFNFPTGHLINPTNNN-TMD-----SSFAGAYKAV 337  
Qy 302 NLASQWNSNYDPTTEIPAPLGTDPFVGKIQGM-----LTQTTRE--DGSTRAHKAT 350  
Db 338 -----SNGIKTDNV---YGLP--VGEITGLPKDPSDIPFNSTTGELVDFSTGKPINN 385  
Qy 351 VSTGSVHFTPKLGSVOYTTDNDFTQNTKFTFVGVIQDGNHONBPQQ----- 401  
Db 386 STAGIVSGRKPGLPPIE--DENGNLFPDSTN--LPI-----DGNOLVNPETNSTVSGSTS 436  
Qy 402 -----WVLPN-----YSGRTGHNH 416  
Db 437 GTTKPKGIPVNGGVVPEEAKQADKQGLIVPPTNSINKDPVTNTQYSNTTGNIN 496  
Qy 417 LAPAVAPTTPGEBQLLFRRSTMPGCGSYPMNLDCLLPQEWVQHFQCEAAPAQSDVALLRF 476  
Db 497 --PETGKVIPTG-----SLPGLNYPSPFN-----TPQQ-----TDEITGKVPDVTVTGLP 537  
Qy 477 VNPDTGEVL 485  
Db 538 YDPSTGEII 546

Search completed: June 1, 2004, 13:57:58  
Job time : 13.8419 secs

Db 774 KELW-----KSRMPVGSQGTPISVSPKTKG 799  
RESULT 15  
US-08-928-361B-30  
; Sequence 30, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSES: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,361B  
; FILING DATE: 12-SEP-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,062  
; FILING DATE: 13-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verny, Hana  
; REGISTRATION NUMBER: 30,518  
; REFERENCE/DOCKET NUMBER: 480.76-1 (HV)  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 650-324-1677  
; TELEFAX: 650-324-1678  
; INFORMATION FOR SEQ ID NO: 30:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1043 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-928-361B-30  
Query Match 3.6%; Score 104.5; DB 3; Length 1043;  
Best Local Similarity 20.5%; Pred. No. 0.86; 203; Indels 227; Gaps 35;  
Matches 125; Conservative 54; Mismatches 54  
Qy 5 SNDANPSDGTANLVPEVNN-----EVMALPEVVGAAIAAPVAGQQNVDPWIRNN 55  
Db 37 SND-EPICGQAGIADTSNLFPPVQTHKSTGLPIDPMVGLPF-DPKSG--NLVHPYTN-- 90  
Qy 56 FVQAFPG-----EFTVSPRNA---PGEILWSAPLGP---DLNPLYSHL-----AR 94  
Db 91 --QTMWGLSVSYLAARNLTVDTDYGLFDITLTGYPLDPVSLFPFNPETGELFDPIDSE 148  
Qy 95 MYNGYAGGFEVQVILAGNAFTAGKIIFAAVPPNF-----PTEG-----LSPSQV 138  
Db 149 INMGTIAGI-VSGISASELSLSQKALIDPATNVVGVGGLLNPATGMIPGFLGPSEQ 207  
Qy 139 TMFPHIIVDVRLQEP-----VLIPLPDVRNNFYHYNQSNDSITKILAMLYTLR 187  
Db 208 TQFSPEIEDGGIIPPEVAANADKFKLSIP-PSVPESIPKDKIDSIDSEL---MYDIES 263  
Qy 188 ANNAGDDVFTVSCRVLTRSP-----DPDFIFLVPTVESRT-KPFTVPIILTVEMSNR 241  
Db 264 GRLIG-----QVSKRPISGIAGDLNPIMKTTQTDSVTGKPID----- 302

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.4562 Seconds  
(without alignments)  
4368.312 Million cell updates/sec

Title: US-09-926-799-5  
Perfect score: 2886  
Sequence: 1 MKMASNDANPSDGSSTANLVP.....VNQFTYLPAMGNGAGRRRAL 539

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/ECT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/ECT05\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10C\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1161.5	40.2	530	14	US-10-314-739-3
2	322	11.2	669	14	US-10-209-507-2
3	300.5	10.4	547	12	US-10-670-695-4
4	300.5	10.4	671	12	US-10-670-695-2
5	297.5	10.3	668	14	US-10-209-507-4
6	135	4.7	6310	12	US-10-282-122A-67793
7	126.5	4.4	1147	15	US-10-327-481A-38
8	122	4.2	26926	9	US-09-759-508B-2
9	116.5	4.0	913	15	US-10-369-493-2294
10	116	4.0	5701	9	US-09-864-761-37319
11	116	4.0	5701	14	US-10-029-386-32438
12	112.5	3.9	710	15	US-10-094-749-2315
13	112	3.9	1344	9	US-09-738-628-6888
14	111.5	3.9	776	9	US-09-765-272-86
15	111.5	3.9	2233	10	US-09-769-787-2

16	109	3.8	607	14	US-10-320-769-4	Sequence 4, Appli
17	109	3.8	782	9	US-09-908-193-47	Sequence 47, Appli
18	109	3.8	782	15	US-10-190-115-39	Sequence 39, Appli
19	109	3.8	782	15	US-10-369-072-39	Sequence 39, Appli
20	109	3.8	3930	12	US-10-282-122A-46817	Sequence 46817, A
21	108.5	3.8	1325	12	US-10-282-122A-44953	Sequence 44953, A
22	108.5	3.8	1767	12	US-10-087-684-40	Sequence 40, Appli
23	108.5	3.8	1767	12	US-10-087-684-41	Sequence 40, Appli
24	108.5	3.8	1767	12	US-10-218-779-40	Sequence 40, Appli
25	108.5	3.8	1767	12	US-10-218-779-41	Sequence 41, Appli
26	107	3.7	1323	14	US-10-156-761-10065	Sequence 10065, A
27	107	3.7	2481	12	US-10-282-122A-43762	Sequence 43762, A
28	106.5	3.7	866	12	US-10-282-122A-70192	Sequence 70192, A
29	106.5	3.7	1119	12	US-10-425-114-53646	Sequence 53646, A
30	106.5	3.7	1143	12	US-10-424-599-237981	Sequence 237981, A
31	106.5	3.7	1439	12	US-10-282-122A-61196	Sequence 61196, A
32	106.5	3.7	2669	12	US-10-016-248-4	Sequence 4, Appli
33	106.5	3.7	3104	12	US-10-016-248-2	Sequence 2, Appli
34	106	3.7	796	9	US-09-801-368-30	Sequence 30, Appli
35	106	3.7	1119	14	US-10-245-802-12	Sequence 12, Appli
36	105.5	3.7	5795	9	US-09-815-242-12610	Sequence 12610, A
37	105.5	3.7	10421	12	US-10-282-122A-61631	Sequence 61631, A
38	105	3.6	925	15	US-10-108-260A-3889	Sequence 3889, Ap
39	105	3.6	1313	12	US-10-282-122A-76863	Sequence 76863, A
40	104.5	3.6	1047	14	US-10-245-802-22	Sequence 22, Appli
41	104.5	3.6	1082	15	US-10-369-493-17890	Sequence 17890, A
42	103.5	3.6	685	15	US-10-104-047-2916	Sequence 2916, Ap
43	102.5	3.6	1794	10	US-09-965-738-299	Sequence 299, App
44	102.5	3.6	1799	10	US-03-965-738-149	Sequence 149, App
45	102.5	3.6	1821	10	US-03-965-738-82	Sequence 82, Appli

RESULT 1  
US-10-314-739-3  
; Sequence 3, Application US/10314739  
; Publication No. US20030129588A1  
; GENERAL INFORMATION:  
; APPLICANT: Estes, Mary K  
; Jiang, Xi  
; Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; Characterize No. US20030129588A1walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Fulbright & Jaworski L.L.P.  
; CITY: Washington, D.C.  
; STATE: <Unknown>  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/10/314,739  
; APPLICATION NUMBER: US/10/314,739  
; FILING DATE: 09-Dec-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX: <Unknown>

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 530 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match          40.2%; Score 1161.5; DB 14; Length 530;
Best Local Similarity 45.4%; Pred. No. 1.3e-98;
Matches 250; Conservative 83; Mismatches 181; Indels 37; Gaps 14;

QY 1 MMWASNDANPS-DGST--ANLVPEVN-NEVMALPEVVGAAIAAPVAGQONVIDPWIRNPF 56
DB 1 MMWASKDATSSVDGASGAGQLPEVNASDPLAMDFVAGSSTAVATAGQNPIDPWIIINF 60

QY 57 VQAPGGEFTVSPNAPGELTWSAPGLDPLNPLYSLHARMYNGYAGGFEVQVILAGNAFTA 116
DB 61 VQAPQGEFTISPNPTPGDVLFDLSLGFHLNPLFLLHLSQMYNGWGNMVRIMLAGNAFTA 120

QY 117 GKIIFAAEPNPFTEGLSPSQVTMPFPHIIVDVRQLEPVLIPLDVRRNPFYHYNQNSDSTI 176
DB 121 GKLIIVCIPPGFSGHNLITIAQATLFPHVIAADVRLDPIEVPLEDVRNVLFNHNRNQOTM 180

QY 177 KLIAMLYPLRANAGDDVFTVSCRVLTRPSDPDFILVLPPTVESRTKPTVPILTVEE 236
DB 181 RLVCMLYPLTRTGGTGDSEFVAVGRVMTCPSPDFNLFELVPPTEVQKTRPFTPLNPLSS 240

QY 237 MNSRRPPILEKLYTGPSSAFVQPNQGRCTTGVLLGTQTLSAVNICTFRGDVTHIAGS 296
DB 241 LNSRAPLPISSGISPDNVQSVQFQNGRCTLDLGRLVGITPVLSLHVAKIRGT-----S 294

QY 297 HDYTMNLASQWNSYNDTEBIPAPLTPDFVPGKIQKMLTQTTREDGSTRAHKATVSTGVS 356
DB 295 NGTVINLTDLGTPFHPFEG-PAPIGPPD-LGGCDWHINMT--QFGHSSQTQVDVDTTPD 350

QY 357 HFTPKLGSVQYTTDNNDFOTGQNTKFTPVGVIQ--DGNHNEPO--QWVLPNYSGRTG 412
DB 351 TFPVHLGSIQ-----ANGISGNY-----GVLSWISPPSPHSGSQVDLWKIPNYGSSIT 400

QY 413 HNVHLAPAVAPTFPGEQLLFFRSTMPGCGYPNNMLDCLLPQEWVQHFQCOEAAPOADVA 472
DB 401 EATHLAPSVYPPGGEVLVFFMSKMPGFGAY--NLPCLLPQEVISHLASEQAPTVEGA 457

QY 473 LLRFVNPDTRVLPECKLHKSQYTV-----AHTGPHDLVTPNGYFRFDSWVNOQFTLAP 528
DB 458 LLHYVDPDTGRNLGEFKAYPDGFLTCVPNGASSGPGQL--PINGVFVFSWVSRRFYQLKP 515

QY 529 MNGAGRRRAL 539
DB 516 VGTASSARGRL 526

RESULT 2
US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-2

Query Match          11.2%; Score 322; DB 14; Length 669;
Best Local Similarity 25.4%; Pred. No. 1.5e-20;
Matches 125; Conservative 69; Mismatches 158; Indels 140; Gaps 22;

QY 11 SDGSTANLVPEVNVEY--MALEPVVGAATAAPVAGQONVIDPWIRNPFVQAPGGEFTV-- 66
DB 125 ADGDSSTITPEQGTLLGVYIAEPESAQMTAAATAATGSKVDSW--ESP-----FSFHT 175

QY 67 -----SPNAPGELTWSAPGLDPLNPLYSLHARMYNGYAGGFEVQVILAGNAFTA 122
DB 176 SVNWSSTETQKLLFKQSLGLPLNPLYLHLSKLYVWWSGVDRFISGSGVFGKLAAL 235

QY 123 AVPPNPTEGLSPSQVT---MFPHIIVDVRQLEPVLIPLDVRRNPFYHYNQNSDSTIKLI 179
DB 236 VPP-----GVDPVQSTSMQLQYPHVLFDAQVPEVIFSPDLKSTLYHLMSDTHIT-SLV 289

QY 180 AMLYTPLRANAGDDVFTVSC--RVLTRPSDPDFILVPP-----TVES----- 222
DB 290 IMVYNDL-INPYANDSNSSCIVTVETKPGDPFKFHLKPPGSMLTGHSIPSDLI 348

QY 223 -----RTKPTVPILTVEEMSRSRFP 243
DB 349 LWIGNRYWSDITDPVIRPFVQANRHFDNQETAGMSTPRFRPITI---TISESNGS--- 402

QY 244 IPLEKLVTPGSSAPVQV--PQNGRCTDGLVLTQTLSAVNICTFRGDVTHIAGSHDYTM 301
DB 403 ----KLGTGVATDYIVPGIPDGWPDITIGBEL-----TPAGDYIITNGSGN--- 444

QY 302 NLASQNSYNDTEBIEIPAPLTPDFVQK-IQGMLTQ-----TTREDGSTRA 346
DB 445 DIATAN--AYDSADVI---TNTNFRGMWICGALQRAWGDKKISSTAFITTAKEGNTLK 499

QY 347 HKATVSTGSHFTPKLGSVQYTTDNNDFOTGQNT-----KFTPVGVIQDGNHNEPOQW 402
DB 500 PSNTIDMTKI-----AVYQDTHVGRDVQTSDDTLAILGYTGIGEQAIQSNRDSVVRI 552

QY 403 VLPNYSGRGTGN 414
DB 553 MLPETGARGGNH 564

US-10-670-695-4
RESULT 3
; Sequence 4, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-4

Query Match          10.4%; Score 300.5; DB 12; Length 547;
```

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 530 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match          40.2%; Score 1161.5; DB 14; Length 530;
Best Local Similarity 45.4%; Pred. No. 1.3e-98;
Matches 250; Conservative 83; Mismatches 181; Indels 37; Gaps 14;

QY 1 MMWASNDANPS-DGST--ANLVPEVN-NEVMALPEVVGAAIAAPVAGQONVIDPWIRNPF 56
DB 1 MMWASKDATSSVDGASGAGQLPEVNASDPLAMDFVAGSSTAVATAGQNPIDPWIIINF 60

QY 57 VQAPGGEFTVSPNAPGELTWSAPGLDPLNPLYSLHARMYNGYAGGFEVQVILAGNAFTA 116
DB 61 VQAPQGEFTISPNPTPGDVLFDLSLGFHLNPLFLLHLSQMYNGWGNMVRIMLAGNAFTA 120

QY 117 GKIIFAAEPNPFTEGLSPSQVTMPFPHIIVDVRQLEPVLIPLDVRRNPFYHYNQNSDSTI 176
DB 121 GKLIIVCIPPGFSGHNLITIAQATLFPHVIAADVRLDPIEVPLEDVRNVLFNHNRNQOTM 180

QY 177 KLIAMLYPLRANAGDDVFTVSCRVLTRPSDPDFILVLPPTVESRTKPTVPILTVEE 236
DB 181 RLVCMLYPLTRTGGTGDSEFVAVGRVMTCPSPDFNLFELVPPTEVQKTRPFTPLNPLSS 240

QY 237 MNSRRPPILEKLYTGPSSAFVQPNQGRCTTGVLLGTQTLSAVNICTFRGDVTHIAGS 296
DB 241 LNSRAPLPISSGISPDNVQSVQFQNGRCTLDLGRLVGITPVLSLHVAKIRGT-----S 294

QY 297 HDYTMNLASQWNSYNDTEBIPAPLTPDFVPGKIQKMLTQTTREDGSTRAHKATVSTGVS 356
DB 295 NGTVINLTDLGTPFHPFEG-PAPIGPPD-LGGCDWHINMT--QFGHSSQTQVDVDTTPD 350

QY 357 HFTPKLGSVQYTTDNNDFOTGQNTKFTPVGVIQ--DGNHNEPO--QWVLPNYSGRTG 412
DB 351 TFPVHLGSIQ-----ANGISGNY-----GVLSWISPPSPHSGSQVDLWKIPNYGSSIT 400

QY 413 HNVHLAPAVAPTFPGEQLLFFRSTMPGCGYPNNMLDCLLPQEWVQHFQCOEAAPOADVA 472
DB 401 EATHLAPSVYPPGGEVLVFFMSKMPGFGAY--NLPCLLPQEVISHLASEQAPTVEGA 457

QY 473 LLRFVNPDTRVLPECKLHKSQYTV-----AHTGPHDLVTPNGYFRFDSWVNOQFTLAP 528
DB 458 LLHYVDPDTGRNLGEFKAYPDGFLTCVPNGASSGPGQL--PINGVFVFSWVSRRFYQLKP 515

QY 529 MNGAGRRRAL 539
DB 516 VGTASSARGRL 526

RESULT 2
US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-2
```

Best Local Similarity 25.3%; Pred. No. 1.1e-18; Matches 141; Conservative 80; Mismatches 238; Indels 99; Gaps 28;

QY 12 DGSIT--APEQGTWGGVIAEPAQMSAQAADWATKSVDSW---EAPFSFHTSVNM 178  
DB 12 DGSIT--APEQGTWGGVIAEPAQMSAQAADWATKSVDSW---EAPFSFHTSVNM 178  
QY 67 SPRNAPGELWSAPGLDNLNLYSHLARMYAGGFEVQVILAGNAFTAGKIIFAAVPP 126  
DB 67 SPRNAPGELWSAPGLDNLNLYSHLARMYAGGFEVQVILAGNAFTAGKIIFAAVPP 126  
QY 55 STSETQKILFKQSLGPLNLYLEHLAKLYVAMSGSIEVRFSISGSGVFGKLAIVVPP 238  
DB 55 STSETQKILFKQSLGPLNLYLEHLAKLYVAMSGSIEVRFSISGSGVFGKLAIVVPP 238  
QY 127 NPFTGLSPSQVT---MPPHIIVDVRQLEPVLILPLDVRNNFYHYNQSDSTIKILAMLY 183  
DB 127 NPFTGLSPSQVT---MPPHIIVDVRQLEPVLILPLDVRNNFYHYNQSDSTIKILAMLY 183  
QY 115 -----GVDPVQSTMLQYPHVLFQARQVEPVIFCLPDLRSTLYHLSMDTDTT-SLVIMVY 292  
DB 115 -----GVDPVQSTMLQYPHVLFQARQVEPVIFCLPDLRSTLYHLSMDTDTT-SLVIMVY 292  
QY 184 TPLRANNAGDDVFTVSC--RVLTSPSPDPDFELVPPTVESRTKPTFVILTVSEMSNR 241  
DB 184 TPLRANNAGDDVFTVSC--RVLTSPSPDPDFELVPPTVESRTKPTFVILTVSEMSNR 241  
QY 169 NDL--INPYANDANSNGCIVTVETKPGDFKPHLLKPP--GSMITHGSPSDLPKTS-- 347  
DB 169 NDL--INPYANDANSNGCIVTVETKPGDFKPHLLKPP--GSMITHGSPSDLPKTS-- 347  
QY 242 FPIPLEKLYTGSSAFVVPQ---QNGR-----CTTDGVLGTTQLSAVNICTFRGDVTTHI 293  
DB 242 FPIPLEKLYTGSSAFVVPQ---QNGR-----CTTDGVLGTTQLSAVNICTFRGDVTTHI 293  
QY 224 --LWIGNRYWSDITDFVIRPFVQANRHFDFNQETAGWSTPRFRPISVTITEQNGAKLGI 405  
DB 224 --LWIGNRYWSDITDFVIRPFVQANRHFDFNQETAGWSTPRFRPISVTITEQNGAKLGI 405  
QY 294 AGSHDYTNLASQNSNYD--PTEBIPAPLGTDPFVGKIQGMLTQTTREDGSTRA--HKAT 350  
DB 294 AGSHDYTNLASQNSNYD--PTEBIPAPLGTDPFVGKIQGMLTQTTREDGSTRA--HKAT 350  
QY 282 GVATDYIVPGIPDGWPDTTIPGELIPA-----GDYA-----ITNGTGNDITATGYDTAD 455  
DB 282 GVATDYIVPGIPDGWPDTTIPGELIPA-----GDYA-----ITNGTGNDITATGYDTAD 455  
QY 351 VSTGSHVF-----TPKLGSVQYTT-----DTNNDFTQNTKFTFVGVIQ 390  
DB 351 VSTGSHVF-----TPKLGSVQYTT-----DTNNDFTQNTKFTFVGVIQ 390  
QY 332 IIKNTNFRGMVYCGSLQRAWGDKKISNTAFITATLDGDNKNKINPCNTIDQSKIIVFQ 437  
DB 332 IIKNTNFRGMVYCGSLQRAWGDKKISNTAFITATLDGDNKNKINPCNTIDQSKIIVFQ 437  
QY 391 D---GNHQNHPQWVLPNYSG---RTGHNHVLAPAVAPTPP-----GEQLLFFRSTM 437  
DB 391 D---GNHQNHPQWVLPNYSG---RTGHNHVLAPAVAPTPP-----GEQLLFFRSTM 437  
QY 392 DNHVKKAKTSDDTLALLGYTGIGEQAIGSDRDRVVRIS--TLPETGARGGNHPIFYKNSI 574  
DB 392 DNHVKKAKTSDDTLALLGYTGIGEQAIGSDRDRVVRIS--TLPETGARGGNHPIFYKNSI 574  
QY 438 PCSGYPNMNLDCLLPQEWVQHFQEEA-----PAQSDVALLRFVNPDTGRVLFCKLH 491  
DB 438 PCSGYPNMNLDCLLPQEWVQHFQEEA-----PAQSDVALLRFVNPDTGRVLFCKLH 491  
QY 451 K--LGVIRSIDVFNQ--ILHTRQLSLNHVLLPPDS--FAVYRII--DSNGSWDFDID 627  
DB 451 K--LGVIRSIDVFNQ--ILHTRQLSLNHVLLPPDS--FAVYRII--DSNGSWDFDID 627  
QY 492 KSGYVTVVHTGPHDLVIP 509  
DB 492 KSGYVTVVHTGPHDLVIP 509  
QY 504 SDGFSFVGVSGFGKLEFP 521  
DB 504 SDGFSFVGVSGFGKLEFP 521

RESULT 4  
US-10-670-695-2  
; Sequence 2, Application US/10670695  
; Publication No. US2004005816A1  
; GENERAL INFORMATION:  
; APPLICANT: Jensen, Wayne A.  
; APPLICANT: Lappin, Michael R.  
; APPLICANT: Rosen, David K.  
; APPLICANT: Andrews, Janet S.  
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE  
; TITLE OF INVENTION: STATUS OF AN ANIMAL  
; FILE REFERENCE: DI-9-1  
; CURRENT APPLICATION NUMBER: US/10/670,695  
; PRIOR FILING DATE: 2003-09-25  
; PRIOR APPLICATION NUMBER: 09/521,738  
; PRIOR FILING DATE: 2000-03-09  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 671  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-10-670-695-2

Query Match 10.4%; Score 300.5; DB 12; Length 671;  
Best Local Similarity 25.3%; Pred. No. 1.5e-18;  
Matches 141; Conservative 80; Mismatches 238; Indels 99; Gaps 28;

QY 12 DGSIT--APEQGTWGGVIAEPAQMSAQAADWATKSVDSW---EAPFSFHTSVNM 178  
DB 12 DGSIT--APEQGTWGGVIAEPAQMSAQAADWATKSVDSW---EAPFSFHTSVNM 178

Best Local Similarity 25.3%; Pred. No. 1.1e-18; Matches 141; Conservative 80; Mismatches 238; Indels 99; Gaps 28;

QY 12 DGSIT--APEQGTWGGVIAEPAQMSAQAADWATKSVDSW---EAPFSFHTSVNM 178  
DB 12 DGSIT--APEQGTWGGVIAEPAQMSAQAADWATKSVDSW---EAPFSFHTSVNM 178  
QY 67 SPRNAPGELWSAPGLDNLNLYSHLARMYAGGFEVQVILAGNAFTAGKIIFAAVPP 126  
DB 67 SPRNAPGELWSAPGLDNLNLYSHLARMYAGGFEVQVILAGNAFTAGKIIFAAVPP 126  
QY 55 STSETQKILFKQSLGPLNLYLEHLAKLYVAMSGSIEVRFSISGSGVFGKLAIVVPP 238  
DB 55 STSETQKILFKQSLGPLNLYLEHLAKLYVAMSGSIEVRFSISGSGVFGKLAIVVPP 238  
QY 127 NPFTGLSPSQVT---MPPHIIVDVRQLEPVLILPLDVRNNFYHYNQSDSTIKILAMLY 183  
DB 127 NPFTGLSPSQVT---MPPHIIVDVRQLEPVLILPLDVRNNFYHYNQSDSTIKILAMLY 183  
QY 115 -----GVDPVQSTMLQYPHVLFQARQVEPVIFCLPDLRSTLYHLSMDTDTT-SLVIMVY 292  
DB 115 -----GVDPVQSTMLQYPHVLFQARQVEPVIFCLPDLRSTLYHLSMDTDTT-SLVIMVY 292  
QY 184 TPLRANNAGDDVFTVSC--RVLTSPSPDPDFELVPPTVESRTKPTFVILTVSEMSNR 241  
DB 184 TPLRANNAGDDVFTVSC--RVLTSPSPDPDFELVPPTVESRTKPTFVILTVSEMSNR 241  
QY 169 NDL--INPYANDANSNGCIVTVETKPGDFKPHLLKPP--GSMITHGSPSDLPKTS-- 347  
DB 169 NDL--INPYANDANSNGCIVTVETKPGDFKPHLLKPP--GSMITHGSPSDLPKTS-- 347  
QY 242 FPIPLEKLYTGSSAFVVPQ---QNGR-----CTTDGVLGTTQLSAVNICTFRGDVTTHI 293  
DB 242 FPIPLEKLYTGSSAFVVPQ---QNGR-----CTTDGVLGTTQLSAVNICTFRGDVTTHI 293  
QY 224 --LWIGNRYWSDITDFVIRPFVQANRHFDFNQETAGWSTPRFRPISVTITEQNGAKLGI 405  
DB 224 --LWIGNRYWSDITDFVIRPFVQANRHFDFNQETAGWSTPRFRPISVTITEQNGAKLGI 405  
QY 294 AGSHDYTNLASQNSNYD--PTEBIPAPLGTDPFVGKIQGMLTQTTREDGSTRA--HKAT 350  
DB 294 AGSHDYTNLASQNSNYD--PTEBIPAPLGTDPFVGKIQGMLTQTTREDGSTRA--HKAT 350  
QY 282 GVATDYIVPGIPDGWPDTTIPGELIPA-----GDYA-----ITNGTGNDITATGYDTAD 455  
DB 282 GVATDYIVPGIPDGWPDTTIPGELIPA-----GDYA-----ITNGTGNDITATGYDTAD 455  
QY 351 VSTGSHVF-----TPKLGSVQYTT-----DTNNDFTQNTKFTFVGVIQ 390  
DB 351 VSTGSHVF-----TPKLGSVQYTT-----DTNNDFTQNTKFTFVGVIQ 390  
QY 332 IIKNTNFRGMVYCGSLQRAWGDKKISNTAFITATLDGDNKNKINPCNTIDQSKIIVFQ 437  
DB 332 IIKNTNFRGMVYCGSLQRAWGDKKISNTAFITATLDGDNKNKINPCNTIDQSKIIVFQ 437  
QY 391 D---GNHQNHPQWVLPNYSG---RTGHNHVLAPAVAPTPP-----GEQLLFFRSTM 437  
DB 391 D---GNHQNHPQWVLPNYSG---RTGHNHVLAPAVAPTPP-----GEQLLFFRSTM 437  
QY 392 DNHVKKAKTSDDTLALLGYTGIGEQAIGSDRDRVVRIS--TLPETGARGGNHPIFYKNSI 574  
DB 392 DNHVKKAKTSDDTLALLGYTGIGEQAIGSDRDRVVRIS--TLPETGARGGNHPIFYKNSI 574  
QY 438 PCSGYPNMNLDCLLPQEWVQHFQEEA-----PAQSDVALLRFVNPDTGRVLFCKLH 491  
DB 438 PCSGYPNMNLDCLLPQEWVQHFQEEA-----PAQSDVALLRFVNPDTGRVLFCKLH 491  
QY 451 K--LGVIRSIDVFNQ--ILHTRQLSLNHVLLPPDS--FAVYRII--DSNGSWDFDID 627  
DB 451 K--LGVIRSIDVFNQ--ILHTRQLSLNHVLLPPDS--FAVYRII--DSNGSWDFDID 627  
QY 492 KSGYVTVVHTGPHDLVIP 509  
DB 492 KSGYVTVVHTGPHDLVIP 509  
QY 504 SDGFSFVGVSGFGKLEFP 521  
DB 504 SDGFSFVGVSGFGKLEFP 521

RESULT 5  
US-10-209-507-4  
; Sequence 4, Application US/10209507  
; Publication No. US20030109033A1  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.2  
; CURRENT APPLICATION NUMBER: US/10/209,507  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: 09/617,594  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: Patent in version 3.0  
; SEQ ID NO 4  
; LENGTH: 668  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-10-209-507-4

Query Match 10.3%; Score 297.5; DB 14; Length 668;  
Best Local Similarity 26.8%; Pred. No. 2.8e-18;  
Matches 110; Conservative 52; Mismatches 147; Indels 101; Gaps 20;

QY 12 DGSIT--APEQGTWGGVIAEPAQMSAQAADWATKSVDSW---EAPFSFHTSVNM 178  
DB 12 DGSIT--APEQGTWGGVIAEPAQMSAQAADWATKSVDSW---EAPFSFHTSVNM 178

```

; ORGANISM: Pseudomonas putida
; US-10-282-122A-67793

Query Match      4.7%; Score 135; DB 12; Length 6310;
Best Local Similarity 23.7%; Pred. No. 0.11;
Matches 108; Conservative 50; Mismatches 189; Indels 108; Gaps 24;

QY 7 DANPS-----DGTANLYPEVNNVMALEPVVGAIAAPVAG---QQNVDPWIRNNFVQ 58
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3575 NASPSATFTADDSTAPAA-----AGLTTPGSAIQSGEAGSTVEVKLADGTLVGVVV 3630

QY 59 AGGEFTV--SPRNAPGEIL-----WSAP---LGPDLNP-----YLSHLARMY 96
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3631 PAGGSTVPLSPAQLDQALNVLTLDAAAGNISQPSQIPAPDITPPALPTDVAVSSDGTAV 3690

QY 97 NGYAGFEVQVILAGNAFTAGKIIIPAAVPPN-----PTEGLSPSQVTMPPHIIVDV-- 148
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3691 TGNAPGASSVTVSDG-----AGNVITVAVNPDSGFSVPLDTPQNNQGVTVV---VVTDAAG 3743

QY 149 RQLEPVLILPLPDVRNNFYHN---QSNDSITKILIAMLYTFLRANNAGDDVFTVSCRVLTR 205
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3744 NDSAPVSVTAPDTTNEPATGLTVSPDGTGCTAEPGSGTVEVRNPDDTV-----RGTVT 3798

QY 206 PSPDFDFILVPPTVESRTKPTVPIITVEEMSNRFPPILEKLYTGSSAFVWQPQNGR 265
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3799 AGPDGTFVIPAPPLASGE---TVDVVVIDPAGNESPEIPL---TGPTGTEVATPSALA 3851

QY 266 CTTDGVLLGTTQLSASVNICTFRGDVTTHIAGS---HDTYTMNLASQN-----WSN 310
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3852 ISVDGFL--TGQGTGSLITVTSGGTILGSAIVGSDGTFRFFQNAQLNAQILQVSAKAT 3910

QY 311 YDPTTEIPAPLGTPDFVGKIQGMLTQTTREDGSTRAH-KATVST-----GSVHFTPKLGSV 365
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3911 VDGQPSVPAIIVAND-----TTAPDAPTQVNLNATGTLTGQGEVGTATVRVTDL 3959

QY 366 Q-----YTTDTNNDFTQNTKFTPVGVIODGNN 394
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3960 QGTLTGATVDSNGLF-----SVSFSP--AVANGQN 3988

RESULT 7
US-10-327-481A-38
; Sequence 38, Application US/10327481A
; Publication No. US20040001864A1
; GENERAL INFORMATION:
; APPLICANT: King, Andrew M.Q.
; APPLICANT: Burman, Alison J.
; APPLICANT: Audonnet, Jean-Christophe F.
; APPLICANT: Lombard, Michel F.A.
; TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease
; FILE REFERENCE: 454313-3178
; CURRENT APPLICATION NUMBER: US/10/327,481A
; PRIOR FILING DATE: 2002-12-20
; CURRENT FILING DATE: PCT/FR01/02042
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: FR 0008437
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 38
; LENGTH: 1147
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
; US-10-327-481A-38

Query Match      4.4%; Score 126.5; DB 15; Length 1147;
Best Local Similarity 22.2%; Pred. No. 0.048;
Matches 78; Conservative 41; Mismatches 144; Indels 89; Gaps 15;

QY 72 PGEILNSAPLGPDLNPNVLSHLARMYNGYAGGFEVQVILAGNAFTAGKIIIPAAVPPNPFTE 131
   ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 160 PFGYLTKELPDTHHGVFGLHVDYSYMRMGMDVEVSAGVNOFNGGCLLVAMVPEWKAFD 219

US-10-282-122A-67793
; Sequence 67793, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 67793
; LENGTH: 6310
; TYPE: PRT
```

QY 132 GLSPQVTFPHIIVDVROLEPVLPLDPVRNMFY-HYNQSNDSSTIKLIAMLYTPIRANN 190  
DB 220 TREKQTLFLPHQFISPRNTMAHTIYVYLGVRNRYDQYKHKPWT--LVVMVLSPLTVSN 277  
QY 191 AGDDVFTVSCRVLTRPSDFDFILVPPFV-----ESRTKPTVPILITVEENSNSRFPPI 245  
DB 278 TA-----APOIKYANIAPYVHVAGELPSKEGIFPVACADGGLVTTDP 323  
QY 246 -----LEKLYTGPSSAFVVPQNGRCTTGDVLLGTQLSAYNIC-TF----- 286  
DB 324 KTADPVYGVKNPPKTYN-----PGRF-----TNLLDVAEACPTFLRPDDGKPYVVT 370  
QY 287 RGDVTHIAGSHDYTNLASQNSVNDPTEIIPALGTPDFVGKIQOMLTQTTRDGSGTRA 346  
DB 371 RADDTRLLAKFD--VSLAKHMSN-----TYLSGIAQYYTQYSGTINL 411  
QY 347 H-KATVSTGSG-----VHFTPKLGSQYTTDT-----NNDFOQTGQNTKET 384  
DB 412 HFMFTGSTDQKARYWVAYIPP--GVETPPDTPEEAHCHIAEWDIGLNSKFT 461  
RESULT 8  
US-09-759-508B-2  
; Sequence 2, Application US/09759508B  
; Publication No. US20020182599A1  
; GENERAL INFORMATION:  
; APPLICANT: Fishman, Mark C.  
; TITLE OF INVENTION: Methods for Diagnosing and Treating Heart Disease  
; FILE REFERENCE: 00786/381002  
; CURRENT APPLICATION NUMBER: US/09/759,508B  
; CURRENT FILING DATE: 2001-01-12  
; PRIOR APPLICATION NUMBER: US 60/175,787  
; PRIOR FILING DATE: 2000-01-12  
; NUMBER OF SEQ ID NOS: 11  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 26926  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-759-508B-2

Query Match 4.2%; Score 122; DB 9; Length 26926;  
Best Local Similarity 19.5%; Pred. No. 17;  
Matches 118; Conservative 83; Mismatches 211; Indels 194; Gaps 30;  
QY 63 EFTVSPRNPAG-----ETLM--SAPLGPDLNPLY-----SHLARMYGVAGGF 103  
DB 14146 EFRVSAENRAGIPSEASDSVLMKDAAYPPGPPSPNPHVDTTKKSASLWKGPHYDGL 14205  
QY 104 EV-----QVILAGNAFTAGKIIFAAVPP-----NFTPEGLSPSQV---TMF 141  
DB 14206 EITGVVVEHQVGDENWIKDTGTALRITQFVVPDLQTKYKNFRISAINDAGVGEPAVI 14265  
QY 142 PHIVDVROLEP-----VLIPL-----PDV-----RNNFYHYNQSN 172  
DB 14266 PDVEIVEREMAPDFELDAELRRLVVRAGLSIRIFVIGRPAPEVTTKONILNKRAN 14325  
QY 173 DSTIKLIAMLYTP-----LRANNAGDDVFTVSCRVLTRPSDFDFILVPP-- 218  
DB 14326 INTESFTLLIIPENRVDYTGKVMTIENPAGKSGFVNVRVLDTPGP---VNLNRPTDI 14382  
QY 219 TVESRTKPTVPILITVEENSNSRFPPIPLEKLYTGPSSAFVVPQNGRCTTDCVLLGTQL 278  
DB 14383 TKDSVTLHWDLPID-----GGSRI-----TNVIVEKR--EATKRSYSTATTK- 14423  
QY 279 SAVNICTFRGDVTHIAGSHDYTNLASQNSVNDPTEIIPALGTPDFVGKIQOMLTQTTRDGSGTRA 329  
DB 14424 --CHKTYK--VTGLSEGCYFFRMAENYGEIETETTEPDKASEAPSP---PDSLNI 14476  
QY 330 IQOMLTQTT-----REDGSTRAHKATVSGSVHFTPKLGSQYTTDTN-----N 373  
DB 14477 MD--ITKSTVSLAWPKPKHDGSKI-----TGVIIEAQRKSGDQWTHITTVKGLECVVR 14528

QY 374 DFQTCQNTKFTPVGVQIQDNNHQB-----POQWVLPNYS-----GRTHNVHL 417  
DB 14529 NLTEGEBYTFQVMVNSAGRSAPRESRPVIVKEQTMLEPDLRGIYQKLVIAKAGDNIKV 14588  
QY 418 APAV-----APTF---PGEQLLFFRSTMPGCGSYPMNMLDCLLPQEWQVHFCQEAAPQSD 470  
DB 14589 EIPVUGRPKPTVTWKKGQIL-----KQTRVNFETTATSTI 14625  
QY 471 VALLRFVNPDTGRVLFECKLHKSQYVTVVAHTGPHDLVIPPNGYFRFDSWVNOFYTLA--P 528  
DB 14626 LNIINECVRSDSGYPPLTARNIVGEVGDVITTIQVHDIPGPPTGPIKDFEVSSDFVTFMSDP 14685  
QY 529 MONGAG 534  
DB 14686 PENDGG 14691  
RESULT 9  
US-10-369-493-2294  
; Sequence 2294, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2294  
; LENGTH: 913  
; TYPE: PRT  
; ORGANISM: Schizosaccharomyces pombe  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(913)  
; OTHER INFORMATION: unsure at all Xaa locations  
US-10-369-493-2294

Query Match 4.0%; Score 116.5; DB 15; Length 913;  
Best Local Similarity 22.4%; Pred. No. 0.29;  
Matches 99; Conservative 48; Mismatches 190; Indels 105; Gaps 22;  
QY 10 PSDGSTANLVPEV-----NNEVMALPEVVGAAIAAPVAGQQNVIDPWIRNNFVQAP-GGE 63  
DB 490 PAVGSTTTSTVEVPPVPLAGNVEVYPTPEPTMVSIVTPVAVGSTTSTVPEV-----VPLAGN 543  
QY 64 FTVSPRNAPGELISAPGLDNLNPLYSHLARMYGVAGGFVQVILAGNAFTAGKIIFAA 123  
DB 544 VEVPTEPETMVSIVTPVAVGSTNSTVPEWPL-----AGNVEV-----NALEPDTMVSIV 593  
QY 124 VPPNPFTGLSPSQVTFPHIIVDVROLEP---VLIPLDPVRNMFYHYNQSNDSSTIKLIA 180  
DB 594 PAVGSTTSTTVEVPPVPLAGN--VEVKLEPETMVSIVPAV-----GSTTSTVPEV- 642  
QY 181 MLYTPLRANNAGDDVFTVSCRVLTRPSDFDFILVPPPTVESRTK--PFTVPILITVEENS 238  
DB 643 ---VPLAGN-----VEVYPTPEPTMVSIVPAVGSTTSTVPEVPPVPLAGNVEV- 686  
QY 239 NSRPPIPLEKLYTGPSSAFVVPQNGRCTTQGV-----LLGTTQLSAYNICTFRQDVTHI 293  
DB 687 -----KVLPEPTMVSIVPAVG-STTSTVPEWPLAGNVEVNALEPDTMVSIVPA 736  
QY 294 AGSHDYTNLASQNSVNDPTEIIPALGTPDFVGKIQOMLTQTTRDGSGTRAHKATVST 353  
DB 737 VGSTTSTV-----PEWVPLAGNVEVYPTPEPTMVSIVTPVAVGST-----T 776

QY 354 GSV-HFTPKLSVQ-YTTDTNDFQ-----TGQNTKFTPVGVIOQGN--NHQNEPQ---Q 401  
Db 777 STVPEVPLAGNVVEYPTETIVSVIPAVGSTTSTVPEVPLAGNVVEYPTETIVS 836  
QY 402 WVLNPNYSGRTGHNHVLAPAVAP 423  
Db 837 VIVP-----AVGSTTSTVPEVVP 854  
RESULT 10  
US-09-864-761-37319  
; Sequence 37319, Application US/09864761  
; Patent No. US20020048763A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; APPLICANT: Chen, Wensheng  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-1  
; CURRENT APPLICATION NUMBER: US 09/864,761  
; CURRENT FILING DATE: 2001-05-23  
; PRIOR APPLICATION NUMBER: US 60/180,312  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: US 60/207,456  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: US 09/632,366  
; PRIOR FILING DATE: 2000-08-03  
; PRIOR APPLICATION NUMBER: GB 24263.6  
; PRIOR FILING DATE: 2000-10-04  
; PRIOR APPLICATION NUMBER: US 60/236,359  
; PRIOR FILING DATE: 2000-09-27  
; PRIOR APPLICATION NUMBER: PCT/US01/00666  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00667  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00664  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00669  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00665  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00668  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00663  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00662  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00661  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: PCT/US01/00670  
; PRIOR FILING DATE: 2001-01-30  
; PRIOR APPLICATION NUMBER: US 60/234,687  
; PRIOR FILING DATE: 2000-09-21  
; PRIOR APPLICATION NUMBER: US 09/608,408  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: US 09/774,203  
; PRIOR FILING DATE: 2001-01-29  
; NUMBER OF SEQ ID NOS: 49117  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 37319  
; LENGTH: 5701  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC010680.3  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 0.95  
; OTHER INFORMATION: EXPRESSED IN HEL100, SIGNAL = 0.94  
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 3.8  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1

OTHER INFORMATION: SWISSPROT HIT: P16419, EVALUATE 2.00e-73  
OTHER INFORMATION: EST\_HUMAN HIT: AW384459.1, EVALUATE 5.00e-90  
OTHER INFORMATION: EST\_HUMAN HIT: AW384459.1, EVALUATE 5.00e-90  
US-09-864-761-37319  
Query Match 4.0%; Score 116; DB 9; Length 5701;  
Best Local Similarity 19.2%; Pred. No. 5.4;  
Matches 111; Conservative 81; Mismatches 203; Indels 182; Gaps 28;  
QY 80 PLGPDLPYL-----SHLARMYNGYAGGPEV-----OVILAGNAFTAGKII 120  
Db 1 PPGPPSPHVTDTTKKSASLAWGKPHYDGLLEITGVVHEHQKVGDEAMIKTTGTALRIT 60  
QY 121 FRAVPP-----NPTEGLSPSQV---TMFPHIIVDRQLEP----- 153  
Db 61 QFVVPDLQTKKXNFSAINDAGVGEPAVDPDVEIVEREMAPDFELDAELRRLTVVRAG 120  
QY 154 ----VLIPL-----PDV-----RNNFHYNOSNDSTIKLAMLVTP-----LRAN 189  
Db 121 LSIRIFVPIKGRPAPEVTWTKONINLKNRANIENTESFTLLIIPECNRYDTGKFVMTIEN 180  
QY 190 NAGDDVFTYSCRVLTSPDFDFIFLVP--TVESRTKPTVPILTVEMSMNSRFPPIPLE 247  
Db 181 PAGKSGFVNRVLDTPGP---VINLRPTDIKDSVTLHWDLPID---GGSRI----- 228  
QY 248 KLYTGPSSAFVQFQNGRCCTDGVLLGTQLSANNICTFGRDVTHTAGSHDYTMNLASON 307  
Db 229 -----TNVIVEKR--EATRKSYSTATTK---CHKCTYK--VTGLSECEYFFRVAEN 274  
QY 308 -MSNYDPE-----EIPAPLGTDFVGVKIQGLMTOTT-----REDGSTRAHKAT 350  
Db 275 EYGIGETETETEPVKASEAPSP---PDSLIMD--ITKSTVSLAWPKPKHDDGSKI--- 325  
QY 351 VSTGSHVHTPKLGSVQYTTDTN-----NDFQTQNTKFTPVGVIOQGNHONE----- 398  
Db 326 --TGVVIEAQRKSDOWTHITTVKGLFCVVRNLTEGEYTFQVMVNSAGRSAPRESRPV 383  
QY 399 --PQMWLPNYS-----GRTHNVHLAPAV-----APTFF---PGEOLLFRSTMPG 439  
Db 384 IVKEQTMPLPELDLRGIYQKLVIAKAGDNIKVEIPVLGRPKPTVTWKKGDIIL----- 435  
QY 440 CSGYPNNWLDCLLQEWVQHFCQEAAPASQSDVALLRFVNPDTGRVLFECKLHKSQYVTV 499  
Db 436 -----KQTORVNFETTATSTILNINECVRSDSGPGPYPLTARNIVEGVGDVI 480  
QY 500 HTGPHDLVIPPNGYFRFDSWVNVQFYTLA--PMNGAG 534  
Db 481 TIQVHDIPGPTGPIKDFEVSSDFVTFSWDPPENDGG 517  
RESULT 11  
US-10-029-386-32438  
; Sequence 32438, Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.  
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; FILE REFERENCE: Aeonica-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 32438  
; LENGTH: 5701  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC010680.2  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.6  
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2.1



```
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.2
; OTHER INFORMATION: SWISSPROT HIT: P16419, EVALUE 2.00e-72
US-10-029-386-32438

Query Match      4.0%; Score 116; DB 14; Length 5701;
Best Local Similarity 19.2%; Pred. No. 5.4;
Matches 111; Conservative 81; Mismatches 203; Indels 182; Gaps 28;

QY 80 PLGPDLPYL-----SHLARMYNGYAGGFV-----QVILAGNAFTAGKII 120
Db 1 PGPSPNPHVTDTTKASGLANGKHYDGLSITGVVHEQKVGDEAWIKOTTGTALRIT 60

QY 121 FAAVPP-----NFTPEGLSPSQV---TMFPHIIVDVRLQLEP----- 153
Db 61 QFVVPDLQTKYKNFRISAINDAGCEPAVIDPVEIVEREMAPDFELDAELRTLVRAG 120

QY 154 -----VLIPL-----PDV-----RNNFHYNOSNDSTIKLIAMLYTP-----LRAN 189
Db 121 LGSIRFVPIKGRPAPEVTTKONILNKRANIENTESFTLLIIPECNRYDTGKFMVTIEN 180

QY 190 NAGDDVFTVSCVRLTRPSDFDFELVPP--TVESRTKPTFTVPILTVEMSNSRPEIPLE 247
Db 181 PAGKSGGFVNVRLDTGPG---VLNLRPTDITKDSVTLHMDLPLID-----GGSRI 228

QY 248 KLYTGPSSAFVQPNQGRCTTGDVLLGTQLSAVNICTPRGDVTIAGSHDVTMNLSON 307
Db 229 -----TNYIIVEKR--EATRSKSYSTATYK--CHKCYK--VTGLSECEYFFRVMAEN 274

QY 308 -WSNYDPT-----EIPAPLGTDFVFGKIQGMLTQTT-----REDGSTRAHKAT 350
Db 275 EYGEIGETTEPVRKASEAPSP---PDSLINMD--ITKSTVSLWPKPKHGGSKI----- 325

QY 351 VSTGSVHFTPKLGSVQYTTDTN-----NDFQGTQNTKFTFPGVIGDGNHNE----- 398
Db 326 --TGVIIEAQRKSGDOWHTITTVKGLCEVVRNLTEGEYTFQVMVNSAGRSAPRESRPV 383

QY 399 ---PQOWLPNYS-----GRTHGNVHLAPV-----APTF---PGEQLLFFRSTWPG 439
Db 384 IVKEQTMPELDRLGIYQKLVITAKAGDNIKVEIPVLGRPKFTVWKKGQDIL----- 435

QY 440 CSGYPNMNLDCLLPQEWQHFQCEAAPAQSDVALLRFVNPDTGRVLFECKLHKSGYVTA 499
Db 436 -----KQQRVNFETTATSTILNINECVRSDSGPVPLTARNIVGEVDVI 480

QY 500 HTGPHDLVLPNGYFRFDSWVNOFTLA--PMNGAG 534
Db 481 TTIQVHDIFGPTGPIKDFEVSSDFVTFSDPPENDGG 517
```

```
RESULT 12
US-10-094-749-2315
; Sequence 2315, Application US/10094749
; Publication No. US20030219741A1
; GENERAL INFORMATION:
; APPLICANT: ISOGAI, TAKAO
; APPLICANT: SUGIYAMA, TOMOYASU
; APPLICANT: OTSUKI, TETSUJI
; APPLICANT: WAKAMATSU, AI
; APPLICANT: SATO, HIROYUKI
; APPLICANT: ISHII, SHIZUKO
; APPLICANT: YAMAMOTO, JUN-ICHI
; APPLICANT: ISONO, YUUKO
; APPLICANT: HIO, YURI
; APPLICANT: OTSUKA, KAORU
; APPLICANT: NAGAI, KEIICHI
; APPLICANT: IRIE, RYOTARO
; APPLICANT: TAMECHIKA, ICHIRO
; APPLICANT: SEKI, NAHIKO
; APPLICANT: YOSHIKAWA, TSUTOMU
; APPLICANT: OTSUKA, MOTOKYU
```

```
; APPLICANT: NAGAHARI, KENJI
; APPLICANT: MASUHO, YASUHIKO
; TITLE OF INVENTION: NOVEL FULL-LENGTH CDNA
; FILE REFERENCE: 084335/0160
; CURRENT APPLICATION NUMBER: US/10/094,749
; CURRENT FILING DATE: 2002-03-12
; PRIOR APPLICATION NUMBER: 60/350,435
; PRIOR FILING DATE: 2002-01-24
; PRIOR APPLICATION NUMBER: JP 2001-328381
; PRIOR FILING DATE: 2001-09-14
; NUMBER OF SEQ ID NOS: 3381
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2315
; LENGTH: 710
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-094-749-2315

Query Match      3.9%; Score 112.5; DB 15; Length 710;
Best Local Similarity 21.3%; Pred. No. 0.45;
Matches 97; Conservative 45; Mismatches 163; Indels 151; Gaps 19;

QY 3 MASNDANPSDGSSTANLVPENNEV--MALEPVVGAIAAPVA---GQONVIDPWIRNFVQ 58
Db 123 MPRFIAAPSTLUSLPKVPQGVTVTMESSIPQASAIPTATISGQGG----- 170

QY 59 APGGFTVSPRNAPGEILWSAPLGPDLNPLYSLHARMYNGVAGGFEVQVILAGNAFTAGK 118
Db 171 HPSNLHHTMTNVQMSIIRSNAPGEPPLHIGASHLPR---GAAA----- 210

QY 119 IIPAAVPPNFTPEGLSPQVTFPHHIIIVDVRLQLEPVLPLDPVRNPNFHYNOSNDSTIKL 178
Db 211 ---AAV-----MSSSKVTT---VLRPTSQLPNAATAQPAVOHIIHQPIQSR----- 250

QY 179 IAMLVTPLRANNGDDVFTVSCRVLTRPSDFDFELVPPTV---ESRTKPTFTVPILTVE 235
Db 251 -----PPVTTGNA-----IPPAVATVSVATRAQSPVITTTA 281

QY 236 EMSNRRFPIPLEKLYTGPSSAFVQ--PQNGRCTTGDVLLGTQLSAVNICTPRGDVTTHIA 294
Db 282 AHATDSALRPTLSIQHPSPAISIORPAQSRDVTTRIL-----PSHPA 325

QY 295 GSHDYTMNLSONWSNYDPTTEIIPALCTDFVFGKIQGMLTQTTREDGSTRAHKATVSTG 354
Db 326 -----LGTPKQQLHTMAQKTIPTGTGTGVAATVAPILATNTIPSTATT-----AG 369

QY 355 SVHFT--PKLGSVQYTTDTNNDFOGTQNTKFTFPGVIGDGNHNEPQOWLPNYSGRGT 412
Db 370 SVSHQAFTSTVTMTVPESHSHATAVTTISNIPVAKV-----VPOO----- 410

QY 413 HNVHLAPAVPTFPGEQQLFFRSTWPGSGY---PN 445
Db 411 -ITHSPRIQPDYPAE-----RSSLIPIGSHRASPEN 440
```

```
RESULT 13
US-09-738-626-6888
; Sequence 6888, Application US/09738626
; Publication No. US20020197605A1
; GENERAL INFORMATION:
; APPLICANT: NAKAGAWA, SATOSHI
; APPLICANT: MIZOGUCHI, HIROSHI
; APPLICANT: ANDO, SEIKO
; APPLICANT: HAYASHI, MIKIRO
; APPLICANT: OCHIAI, KEIKO
; APPLICANT: YOKOI, HARUHIKO
; APPLICANT: TATEISHI, NAOKO
; APPLICANT: SENOH, AKIHIRO
; APPLICANT: IKEDA, MASATO
; APPLICANT: OKAZAKI, AKIO
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-125
; CURRENT APPLICATION NUMBER: US/09/738,626
```

Wed Jun 2 09:13:30 2004

ADDRESS: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLY APPLICATION NUMBER: US/09/765,272
FILING DATE: 22-Jan-2001
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLY APPLICATION NUMBER: 08/961,083
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 86:
SEQUENCE CHARACTERISTICS:
LENGTH: 776 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 86:
US-09-765-272-86
Query Match 3.9%; Score 111.5; DB 9; Length 776;
Best Local Similarity 19.9%; Pred. No. 0.65;
Matches 111; Conservative 63; Mismatches 235; Indels 149; Gaps 25;
QY 5 SNDANPSDGSSTANLVPEVNEVMALEPVVGAATAAPVAGQONVIDPWIRNNFVQAGGEF 64
DB 62 ASDSNPSD-----PVS-----NVNDKLIISYN--NQPANRW 89
QY 65 T-----VSPRNPAGILMSAPLGPDLNPLYSLHARMYNGYAGGFEVQVILAGNAFTAKII 120
DB 90 TNWARTNPEASGVGLFGDSGI-----LS--KASVDNLSVGFHED--HGVGVKPSYVI 137
QY 121 PAAVPPNPFTBGLSPSQVTMPFPHIIVDVRLQLEPV-----IPLPD 160
DB 138 EYVVGKTVTPAKNPSFVGNEDHVFNDSANWKPVTLNKAQAKAGEMNHFSFDKVTYA 197
QY 161 VRNNFYHYNQSDSTIKLIAMLYTPLRANNAGDDVFTVSCRVLTRPSPDFDFIFLVPPTV 220
DB 198 VRIAMVKAADNKRKGTSTIEVQIFAKVAAAKOGQTRIQVDGKDLANFNPDLDYDL--ESV 255
QY 221 ESRTKPFVTPILTVEEMSNRPPIPLEKLYTGPSSAFVVPQNG-----RCTTDGLV 272
DB 256 DGR-----VPAVTA--SVNNGLATVVPVREGEPRVIAKAENGIDILGEYRLHFTKDKSL 309
QY 273 LGTTQLSAYNICFTR--GDV-----THIAGSHDYTMNLASQWNSVNDPTEEPAPLG 322
DB 310 LSHKPVAAVKQARLLQVGOALELTPKVPVYFTGDKGYETKDLTVEW-----EVPANL 363
QY 323 TPDFVKGIOGMLTQTTREDGSTRAHKATVSTGSHFTPKLGSV-----QYTTDNDPQT 377
DB 364 TK-----AGQFTVRGRVGLGSNLVAEIT-----VRVTDKLGFTLSDNPNYDENSNAQAFAS 412
QY 378 GQN-----TKFTFPVGVIQDGNHNEPQQWLVNPNYSGRTCHNVHLAPAVAPTTPGQLLFF 433
DB 413 ATNDIDKNSHDRVVDYLDNDGDHSEN--RRWT--NWSPTSSNPEVSAGV-----IF 458
QY 434 RSTMWPGCSGYFNWNLDCLLPQEWVQHFCEAAAPQASDVALLRFVNPDTGRVLFECKLHKS 493

CURRENT FILING DATE: 2000-12-18
PRIOR APPLICATION NUMBER: JP 99/377484
PRIOR FILING DATE: 1999-12-16
PRIOR APPLICATION NUMBER: JP 00/159162
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: JP 00/280988
PRIOR FILING DATE: 2000-08-03
NUMBER OF SEQ ID NOS: 7059
SOFTWARE: Patentin ver. 3.0
SEQ ID NO 6888
LENGTH: 1344
TYPE: PRT
ORGANISM: Corynebacterium glutamicum
US-09-738-626-6888
Query Match 3.9%; Score 112; DB 9; Length 1344;
Best Local Similarity 21.8%; Pred. No. 1.4; Mismatches 178; Indels 250; Gaps 34;
Matches 132; Conservative 46;
QY 1 MKWASNDANPSDGS--TAN-LVPE-----VNEVMALEPVVGAATAAPVAGQONVIDPWI 52
DB 158 VEITNTTAPIDFSDYTLNLYPQDEFTTNEAVAAEP--GDVLIQP--GKSLVF--WI 211
QY 53 RNNFVQAGGEFTVSPRNPAGELMSAPLGPDLNPLYSLHARMYNGYAGGFEVQVILAGN 112
DB 212 KN-----GPDEATAADFNA--EYGTNLEAGKDLVEISS--GGWANGTARQMGIQT-NTGH 262
QY 113 AFTAGKIIFAAVPPNPFTGSLSPSQVTMPFPHIIVDVRLQ-LEPVLI-----PLPDVRNNFYH 167
DB 263 IVNRGFYNMAGADVKANEGL-----HFAVDESLLKQTLVSGGAPTEGT-----VY 309
QY 168 YNQSDNSTIKLIAMLYTPLRANNAGDDVFTVSCRVLTRPSPDFDFIFLVPPTVESRTKPF 227
DB 310 TSQIPNPLSAVIADSSVPLITDNTATSI-----NPAEPFTFAPNI----- 349
QY 228 TVPILTVEMSNRFPPIPLEKLYTGPSSAFVVPQNGRCCTDGVLLGTTQLSAYNICFTR 287
DB 350 -----TDDVQVRTATL----- 360
QY 288 GDVTHIAGSHDYTMNLASQWNSNYDPTTEIPAPLGPDPFGVKIQGMLTQTTREDGSTRA 346
DB 361 HVTSSAGEAATTNL-----TEDDGSFNWA 385
QY 347 HKATVSTGSHFTPKLGSVQYTTDNDNFQGTQNTKFTPVGVIOGNNHQN-----EPQQ 401
DB 386 LPAADLTGKSWF-----EYTVTATDGF-----NSVTTEPRVTVDGANTDPLRLNLENQ 435
QY 402 WVLPNYSGRT-----GHNVL-----APAV-----APTTPGSQL--LFFRSTMP 438
DB 436 WV-----SGTVDVIGASDVFGDKLELLIDDAPAVTNSSLASAAPTFAEMVTQTDVFRNGI- 490
QY 439 GCSGYPNM-----NLDCLLPQEWVQHFCE-----AAPA-----QS 469
DB 491 -LAGGEELRIPOQTYANTETISTPVLPHYNEDGTLTVSVYAGTKAAPEIDLNENDDF 549
QY 470 DVALLRFVNPDTGRVLFECKLHKSQVTVVAHTGPHDLVIPNGYFRFDSWVNFYTLAPM 529
DB 550 QIRNLRLLIPD-GRTL-----TPAGISDSNAWLN-----M 578
QY 530 GNGAGR 535
DB 579 GDSAGK 584

RESULT 14
US-09-765-272-86
Sequence 86, Application US/09765272
Patent No. US20020061545A1
GENERAL INFORMATION:
APPLICANT: Choi et. al.
TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines
NUMBER OF SEQUENCES: 452
CORRESPONDENCE ADDRESS:

Db 459 REN-----GKIVERTVQGVQFFADSGTDAPSKVLERYVGPE-----FEVPTYYS 505

QY 494 GYVTVAHGTGPHDLVIPPN 511

Db 506 NY--QAYDADHPFNPN 521

RESULT 15

US-09-769-787-2  
; Sequence 2, Application US/09769787  
; Publication No. US20030091577A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbial Technics Limited  
; APPLICANT: Gilbert, Christophe FG  
; APPLICANT: Hansbro, Philip M  
; TITLE OF INVENTION: Proteins  
; FILE REFERENCE: PWC/P21129WO  
; CURRENT APPLICATION NUMBER: US/09/769,787  
; PRIOR FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: GB 9816337.1  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: US 60/125164  
; PRIOR FILING DATE: 1999-03-19  
; NUMBER OF SEQ ID NOS: 388  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 2233  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-769-787-2

Query Match 3.9%; Score 111.5; DB 10; Length 2233;  
Best Local Similarity 19.9%; Pred. No. 3.3;  
Matches 111; Conservative 63; Mismatches 235; Indels 149; Gaps 25;

QY 5 SNDANPSDGSANLPEVNVNEWVALEPVVGAIAAPVAGQQNVDPWIRNPFVQAPGGEF 64  
Db 1484 ASDSNPSD-----PVS---NVNDKLISYN--NQPANRW 1511  
QY 65 T-----VSPRNAPCEILWSAPLPGDLPNLPYLHLARMYNGVAGGFEVQVILAGNAFTAGKII 120  
Db 1512 TNNRTNPEASGVLFQDSGI-----LS--KRSVDNLISVGFHED---HGVGVPKSYVI 1559  
QY 121 FAAPVPNPFTEGLSPSQVTMFPHIIVDVRLQLEPVL-----IPLPD 160  
Db 1560 EYVVGKTVPTAPKNPSFVGNEDHVFENDSANWKPVTNLKAPQKAGEMNHFSFQKVETVA 1619  
QY 161 VRNPFHYNQSDSTIKLIAMLYTPLRANNAGDDVFTVSCRVLTRPSPDFEFILVPPTV 220  
Db 1620 VRIRMVKADNKRGT SITEVQIFAKQVAAAKQSGQTRIQVDGKDLANFNPDLDYIL--BSV 1677  
QY 221 ESRTKPFVTPILTVEMSNRFPIPLEKLYTGPSSAFVVPONG-----RCTTDGVL 272  
Db 1678 DGK-----VPATVA-SVSNGLATVVPVSREGEVPVRVIAKENGDIILGEYRLHFTKDKSL 1731  
QY 273 LGTTQLSAYNICFTR--GDV-----THIAGSHDYTMNLASQNSWNSYDPTTEIPAPLG 322  
Db 1732 LSHKPVAVKQARLLQVQALELPTKVPPVFTGKGYETKDLTVEM-----EEVPAENL 1785  
QY 323 TPDFVGKIQGMLTQITREDCGSTRAHKATVSTGSVHFTPKLSV-----OYTTDNDPOT 377  
Db 1786 TK-----AGQFTVRGRVLGSNLVABIT-----VRVTDKLGTELSDNPYDENSNOQAFAS 1834  
QY 378 GQN-----TKFTPVGVITQDGNHNEPQOWVLNPGYSGRTGHNHVLAPAVAPTFFGQQLPF 433  
Db 1835 ATNDIDKNSHDRVDYLNDDGHSN--RRWT--NWSPTSSNPEVSAGV-----IF 1880  
QY 434 RSTMPGCSGYNNMLDCLLPQEWVQHFCQEAAPQSDVALLRFVNPDTGRVLFECKLHKS 493  
Db 1881 REN-----GKIVERTVQGVQFFADSGTDAPSKVLERYVGPE-----FEVPTYYS 1927  
QY 494 GYVTVAHGTGPHDLVIPPN 511

Db 1928 NY--QAYDADHPFNPN 1943

Search completed: June 1, 2004, 14:04:36  
Job time : 37.4562 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 47.0765 Seconds  
(without alignments)  
3289.030 Million cell updates/sec

Title: US-09-926-799-6  
Perfect score: 2896  
Sequence: 1 MKMASNDAPNSDGAAGLVP.....VNPFYTLAPMTGNGRRRIQ 548

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_25Jan04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2896	100.0	548	4	AAB49705 Small rou
2	2813	97.1	548	5	AAU91272 Norwalk v
3	2124	73.3	550	4	AAB49709 Small rou
4	2117.5	73.1	535	4	AAB49707 Small rou
5	2018	69.7	540	4	AAB49706 Small rou
6	1947.5	67.2	539	4	AAB49704 Small rou
7	1931	66.7	542	4	AAB49708 Small rou
8	1678.5	58.0	541	4	AAB49710 Small rou
9	1211	41.8	530	2	AAR50972 Norwalk v
10	1211	41.8	530	7	ADC72176 Norwalk v
11	1209	41.7	530	4	AAB49701 Small rou
12	1178.5	40.7	545	4	AAB49700 Small rou
13	1164	40.2	530	2	AAR57091 Small rou
14	1143.5	39.5	544	4	AAB49703 Small rou
15	1124.5	38.8	546	4	AAB49702 Small rou
16	285.5	9.9	622	4	AAB47045 Feline ca
17	281	9.7	623	4	AAB47044 Feline ca
18	279.5	9.7	668	4	AAB67462 Amino aci
19	277.5	9.6	579	2	AAW08143 RHDV caps
20	276	9.5	623	4	AAB47043 Feline ca
21	275.5	9.5	547	4	AAM50108 Feline ca
22	275.5	9.5	671	4	AAM50107 Feline ca
23	262	9.0	668	2	AAR10686 Feline ca
24	262	9.0	668	4	AAE04304 Feline ca
25	262	9.0	669	4	AAB67461 Amino aci

ALIGNMENTS

RESULT 1  
AAB49705  
ID AAB49705 standard; protein; 548 AA.  
XX  
AC AAB49705;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 6.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX  
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.

Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
WPI: 2001-080948/09.  
DR N-PSDB; AAF29146.  
XX  
PT Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.  
XX  
PS Claim 1; Page 52-54; 84pp; Japanese.  
CC This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks  
XX  
SQ Sequence 548 AA;  
Query Match 100.0%; Score 2896; DB 4; Length 548;  
Best Local Similarity 100.0%; Pred. No. 2.6e-247;

26	213	7.4	40	5	AAU91274	Norwalk v
27	197	6.8	40	5	AAU91273	Norwalk v
28	194	6.7	40	5	AAU91275	Norwalk v
29	128.5	4.4	2137	5	ABP39618	Staphyloc
30	122	4.2	2206	2	AAR22210	True type
31	121.5	4.2	522	6	ABP56883	Staphyloc
32	121.5	4.2	1831	6	ABU43109	Protein e
33	121.5	4.2	1870	6	ABU19019	Pathogen
34	121	4.2	338	7	ADC06674	Nervous n
35	121	4.2	776	2	AAW55105	Streptoco
36	121	4.2	776	5	ABP54599	S. pneumo
37	121	4.2	776	7	ADC45167	S. pneumo
38	121	4.2	2233	6	ABU01029	S. pneumo
39	121	4.2	2234	3	AAH81502	Streptoco
40	119	4.1	934	1	AAH20016	Sequence
41	118.5	4.1	1541	4	ABG29253	Novel hum
42	116.5	4.0	1147	5	ABB76724	Foot and
43	116	4.0	832	4	ABG23820	Novel hum
44	115	4.0	10431	6	ABU54861	Human CAL
45	114.5	4.0	3892	6	ADA34216	Acinetoba

	Matches	548;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
Qy	1	MKMASNDAA	PSNDCAGL	VP	EINNEAMALDPVAGAAIAAPLTGQONIIDPIMMNFVQAP	60				
Dd	1	MKMASNDAA	PSNDCAGL	VP	EINNEAMALDPVAGAAIAAPLTGQONIIDPIMMNFVQAP	60				
Qy	61	GGFTVS	PRNSPGEVLNLNLELGPEINPYLAHLARMYNGVAGGFVEQVVLAGNAFTAGKII	120						
Dd	61	GGFTVS	PRNSPGEVLNLNLELGPEINPYLAHLARMYNGVAGGFVEQVVLAGNAFTAGKII	120						
Qy	121	FAAIPPNFPI	DNLSAAQITWCPHVI	VVDVROQL	EPVNLPMPDVNNPFHYNOGSDSRLRIA	180				
Dd	121	FAAIPPNFPI	DNLSAAQITWCPHVI	VVDVROQL	EPVNLPMPDVNNPFHYNOGSDSRLRIA	180				
Qy	181	MLYTPLRANSGDDVF	TSCRVLTRSPDPSFNFLVPPTVESKTPFTLPILITISEMSNS	240						
Dd	181	MLYTPLRANSGDDVF	TSCRVLTRSPDPSFNFLVPPTVESKTPFTLPILITISEMSNS	240						
Qy	241	RFPVPIESLHTSPTENIV	OQCGRVTLDGELMGTTQLLPSCAPRGVLTSTSRASDQ	300						
Dd	241	RFPVPIESLHTSPTENIV	OQCGRVTLDGELMGTTQLLPSCAPRGVLTSTSRASDQ	300						
Qy	301	ADTATPRLFNYWHVQDLNLTGPYDPAEDIPGLGTDP	GRGVFGVASQRNLDSSTTRAH	360						
Dd	301	ADTATPRLFNYWHVQDLNLTGPYDPAEDIPGLGTDP	GRGVFGVASQRNLDSSTTRAH	360						
Qy	361	EAKVDTTAGRTFPK	LGSLEISTDSDDFDQNPQTKFTPVGTGVONEAE	FQWSLPDYSGOF	420					
Dd	361	EAKVDTTAGRTFPK	LGSLEISTDSDDFDQNPQTKFTPVGTGVONEAE	FQWSLPDYSGOF	420					
Qy	421	THNMOLA	PAPA	NPPGEOLLEFRSQLPSSGSRSGNVLDCLVPQBWQHFFQESAPAQTV	480					
Dd	421	THNMOLA	PAPA	NPPGEOLLEFRSQLPSSGSRSGNVLDCLVPQBWQHFFQESAPAQTV	480					
Qy	481	ALVRVYNPD	TGKVLF	FEAKHLKLGFMTIANNQDSPITVPPNGYFRPFESVWNPFYTLPMGT	540					
Dd	481	ALVRVYNPD	TGKVLF	FEAKHLKLGFMTIANNQDSPITVPPNGYFRPFESVWNPFYTLPMGT	540					
Qy	541	GNRRRIQ	548							
Dd	541	GNRRRIQ	548							

RESULT 2	
AAU91272	
ID	AAU91272 standard; protein; 548 AA.
XX	
XX	
AC	AAU91272;
XX	
DT	18-JUN-2002 (first entry)
XX	
DE	Norwalk virus associated polynucleotide #1.
XX	
XX	Nowalk virus; monoclonal antibody; geno group I; geno group II;
KW	immunological detection; food; viral infection.
KW	
XX	
OS	Norwalk virus.
XX	
PN	JP2002020399-A.
XX	
PD	23-JAN-2002.
XX	
PF	10-JUL-2000; 2000JP-00208151.
XX	
PR	10-JUL-2000; 2000JP-00208151.
XX	
PA	(OSAP ) OSAKA PREFECTURE.
PA	(NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.
PA	(IATR ) IATRON LAB INC.
XX	
DR	WPI; 2002-287412/33.
XX	
PT	A monoclonal antibody useful in the immunological detection and diagnosis

	of Norwalk virus infection.
PT	Disclosure; Page 12-13; 24pp; Japanese.
XX	The invention describes a monoclonal antibody recognising Norwalk virus,
XX	a capsid protein of Norwalk virus, or a common antigen epitope on the
CC	capsid protein molecule of geno group I and geno group II. The antibody
CC	is useful for immunological detection and quantitative analysis of
CC	Norwalk virus in foods and the serum of infected patients. This sequence
CC	represents a Norwalk virus associated protein described in the invention
XX	Sequence 548 AA;
SQ	Query Match           97.1%; Score 2813; DB 5; Length 548; Best Local Similarity   97.8%; Pred. No. 5.8e-240; Matches 532; Conservative   6; Mismatches   6; Indels     0; Gaps     0;
QY	1 MKMASNDAAPSNDGAAGLVPPEINNEAMALDPVAGAIAAPLTGQQNIIDPMWNFVOAP 60   DB 1 MKMASSSRAAPSNDGAAGLVEPEINNEAMALDPVAGAIAAPLTGQQNIIDPMWNFVOAP 60 
QY	61 GGFEFTVSPPRSFGVLNLIELGPEINPYLAHLARMYNGYAGGFVVQLAGNAFTAGKII 120 
DB	61 GGFEFTVSPPRSFGVELNLIELGPEINPYLAHLARMYNGYAGGFVVQLAGNAFTAGKII 120 
QY	121 FAAIPPNPFDNLSAAGITWCPHVI DVVRQLEPVNLMPEDVNPFHHYNQGSDSLRLRIA 180 
DB	121 FAAPIPNPFDNLSAAGITWCPHVI DVVRQLEPVNLMPEDVNPFHHYNQGSDSLRLRIA 180 
QY	181 MLYTPLRANSGDDVFVTSCRVLTRSPDPSFNFLVPPTVESKTFFTLPILTISEMSNS 240 
DB	181 MLYTPLRANSGDDVFVTSCRVLTRSPDPSFNFLVPPTVESKTFFTLPILTISEMSNS 240 
QY	241 RFPVPIESLHTSPTENIVVQCNGRVTLDGELMGTQTLLPSQICAFRGVLTSTSRASDQ 300 
DB	241 RFPVPIESLHTSPTENIVVQCNGRVTLDGELMGTQTLLPSRICAFGRVLTSTSRASDQ 300 
QY	301 ADATAPRLFNYWHVDLNLMNGTYPDPAEDI PGPLGTDPGRGVFGVASQRNLDDSTTRAH 360 
DB	301 ADIATPRLFNYWHVQLDLNMNGTYPDPAEDI PGPLGTDPGRGVFGVASORNPDSSTTRAH 360 
QY	361 EAKVDTTAGRTFTPKLGSLSEISTSDDDFDQNQPKFTTPVGIGDVNEAEFOQWSLPDYSGOF 420 
DB	361 EAKVDTTAGRTFTPKLGSLSEISTESDFDNQOPTRFTFPGVIGDVNEADFQOWSLPDYSOGF 420 
QY	421 THNNILAPAVANFGEOLLTFRRSQLPSSGGRSNGVLDCLVPOEWWQHFFYESAPAQTQV 480 
DB	421 THNNILAPAVANFGEQLLTFRRSQLPSSGGRSNGVLDCLVPOEWVQHFFYESAPAQTQV 480 
QY	481 ALVRYVNPDTGKVLFEAKLHKLGFMFTIANNGDSPITVPPNGYFRFESWVNFPTLPMGT 540 
DB	481 ALVRYVNPDTGTVLFEAKLHKLGFMFTIAKNGDSPITVPPNGYFRFESWVNFPTLPMGT 540 
QY	541 GNCR 544 
DB	541 GNCR 544 

RESULT 3	
AAB49709	
ID	AAB49709 standard; protein; 550 AA.
XX	
XX	
AC	AAB49709;
XX	
DT	04-APR-2001 (first entry)
XX	
DE	Small round structured virus protein SEQ ID 10.
XX	
KW	Small round structured virus; SRSV; food poisoning.
XX	
OS	Small round structured virus.
XX	
PN	WO200079280-A1.

DR  
XX  
XX  
PT

XX PD 28-DEC-2000.  
XX PF 22-JUN-2000; 2000WO-JP004095.  
XX PR 22-JUN-1999; 99JP-00175928.  
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX PA (DENK-) DENKA SEIKEN KK.  
XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX DR WPI; 2001-080848/09.  
XX DR N-PSDB; AAF29150.  
XX XX  
XX PT Kit for the detection and typing of small round-structured virus (SRSV)  
XX PT strains for investigation of food poisoning outbreaks, contains  
XX PT antibodies.  
XX XX  
XX PS Claim 1; Page 62-64; 84pp; Japanese.  
XX XX  
XX CC This invention relates to a kit for the detection and typing of small  
XX CC round structured virus (SRSV) strains. The kit contains antibodies  
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,  
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
XX CC used for detecting and typing strains of SRSV in order to prevent the  
XX CC spread of infection and to examine the epidemiology of outbreaks  
XX XX  
XX SQ Sequence 550 AA;  
  
Query Match 73.3%; Score 2124; DB 4; Length 550;  
Best Local Similarity 71.4%; Pred. No. 6e-179;  
Matches 396; Conservative 72; Mismatches 75; Indels 12; Gaps 4;  
  
QY 1 MKMASNDAPSDNDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNNIIDPWIMNFVQAP 60  
DB 1 MKMASNDAPSDNDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNNIIDPWIMNFVQAP 60  
  
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGFEVQVVLGNAFTAGKII 120  
DB 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGFEVQVVLGNAFTAGKII 120  
  
QY 121 FFAIPNPFIDNLSAAQITMCPHVIVDVRQLEPVLNLPMPDVNRNFFHYNQGSDSLRLIA 180  
DB 121 FFAIPNPFIDNLSAAQITMCPHVIVDVRQLEPVLNLPMPDVNRNFFHYNQGSDSLRLIA 180  
  
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSPDFSNFLVPPVTSKTKPFTLPILTISEMNS 240  
DB 181 MLYTPLRANSGDDVFTVSCRVLTRPSPDFSNFLVPPVTSKTKPFTLPILTISEMNS 240  
  
QY 241 RPPVPISLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPQICAFRGLVTRSTGRASDQ 300  
DB 241 RPPVPISLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPQICAFRGLVTRSTGRASDQ 300  
  
QY 301 ADTATPRLFNYHVLQDNLNGTTPYDPAEDIPGLTDPDFRGKVFQVSGQRNLS - - - - 355  
DB 301 ADTATPRLFNYHVLQDNLNGTTPYDPAEDIPGLTDPDFRGKVFQVSGQRNLS - - - - 355  
  
QY 356 --TTAHEAKVDTTAGRTFKLSLEISTDDDFONQTKTPVIGVIGVNEAEFCQWSL 413  
DB 356 --TTAHEAKVDTTAGRTFKLSLEISTDDDFONQTKTPVIGVIGVNEAEFCQWSL 413  
  
QY 359 HYATRAHEAHIDTDPKAPKLGTLIKSGSDDFNTNPIRTFVPMGDN - - - - -WQWEL 415  
DB 359 HYATRAHEAHIDTDPKAPKLGTLIKSGSDDFNTNPIRTFVPMGDN - - - - -WQWEL 415  
  
QY 414 PDYSGQFTHMNLAPAVAFNPFGEOLLRPSOLPSSGGRSNGVLVCLVQEWQHFYQES 473  
DB 414 PDYSGQFTHMNLAPAVAFNPFGEOLLRPSOLPSSGGRSNGVLVCLVQEWQHFYQES 473  
  
QY 474 APAQTOVALRVYVNDGTGVLFEAKLHKLGMFTIANNGSDPTTPPNPGYFRPESWVNPY 533  
DB 474 APAQTOVALRVYVNDGTGVLFEAKLHKLGMFTIANNGSDPTTPPNPGYFRPESWVNPY 533  
  
QY 476 AFSQSAVALRVYVNDGTGVLFEAKLHKLGMFTIANNGSDPTTPPNPGYFRPESWVNPY 533  
DB 476 AFSQSAVALRVYVNDGTGVLFEAKLHKLGMFTIANNGSDPTTPPNPGYFRPESWVNPY 533  
  
QY 534 TLAPMGWGRRRIQ 548  
DB 534 TLAPMGWGRRRIQ 548

DB 536 TLAPMGWGRRRIQ 550  
  
RESULT 4  
AAB49707  
ID AAB49707 standard; protein; 535 AA.  
XX AC AAB49707;  
XX DT 04-APR-2001 (first entry)  
XX DE Small round structured virus protein SEQ ID 8.  
XX KW Small round structured virus; SRSV; food poisoning.  
XX OS Small round structured virus.  
XX PN WO200079280-A1.  
XX PD 28-DEC-2000.  
XX PF 22-JUN-2000; 2000WO-JP004095.  
XX PR 22-JUN-1999; 99JP-00175928.  
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX PA (DENK-) DENKA SEIKEN KK.  
XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX DR WPI; 2001-080848/09.  
XX DR N-PSDB; AAF29148.  
XX PT Kit for the detection and typing of small round-structured virus (SRSV)  
XX PT strains for investigation of food poisoning outbreaks, contains  
XX PT antibodies.  
XX PS Claim 1; Page 57-59; 84pp; Japanese.  
XX XX  
XX CC This invention relates to a kit for the detection and typing of small  
XX CC round structured virus (SRSV) strains. The kit contains antibodies  
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,  
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
XX CC used for detecting and typing strains of SRSV in order to prevent the  
XX CC spread of infection and to examine the epidemiology of outbreaks  
XX XX  
XX SQ Sequence 535 AA;  
  
Query Match 73.1%; Score 2117.5; DB 4; Length 535;  
Best Local Similarity 70.5%; Pred. No. 2.2e-178;  
Matches 387; Conservative 74; Mismatches 73; Indels 15; Gaps 4;  
  
QY 1 MKMASNDAPSDNDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNNIIDPWIMNFVQAP 60  
DB 1 MKMASNDAPSDNDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNNIIDPWIMNFVQAP 60  
  
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGFEVQVVLGNAFTAGKII 120  
DB 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGFEVQVVLGNAFTAGKII 120  
  
QY 121 FFAIPNPFIDNLSAAQITMCPHVIVDVRQLEPVLNLPMPDVNRNFFHYNQGSDSLRLIA 180  
DB 121 FFAIPNPFIDNLSAAQITMCPHVIVDVRQLEPVLNLPMPDVNRNFFHYNQGSDSLRLIA 180  
  
QY 181 MLYTPLRANSGDDVFTVSCRVLTRPSPDFSNFLVPPVTSKTKPFTLPILTISEMNS 240  
DB 181 MLYTPLRANSGDDVFTVSCRVLTRPSPDFSNFLVPPVTSKTKPFTLPILTISEMNS 240  
  
QY 241 RPPVPISLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPQICAFRGLVTRSTGRASDQ 300  
DB 241 RPPVPISLHTSPTEINIVVQCONGRVTLTGELMGTTQLLPQICAFRGLVTRSTGRASDQ 300  
  
QY 534 TLAPMGWGRRRIQ 548  
DB 534 TLAPMGWGRRRIQ 548

QY 301 ADTATPRLFNYYVHVLQNLNGTPYDPAEDIPGPIGTDPDRGKVFVGASQRNLDSTTRAH 360  
Db 298 -----NHWNQVNTNINGTPDPTGDPAPLGTDFSGKLFVLSQRDHDAACKSH 348  
QY 361 EAKVDTTAGRTFPTKLGSLLEIST-DSDDFDQNPQTKFTPTVGVIGVNDNEAFQOWSLPDYSGQ 419  
Db 349 DAVIATNSAKFTPKLGAIGTWEEDDVHINQPTKFTPVGL-FENEG-FNQTWLTNYSGA 406  
QY 420 FTHNMNLAPAVAPNPFGEQQLLFFRSQSSGRSNGVLDCLVPOEWVQHEYESAPAQTO 479  
Db 407 LTLNMGGLAPPVAPTPEGEQILFFRSHIPLKGGVADPVICLLPQEWIQHLYQESAPQSOD 466  
QY 480 VALVRYVNPDTGKVLFEAKHLKGFMTIANNGDSPTVPPNGYFRFESWVAPFYTLAPMG 539  
Db 467 VALIRFTNPTGKVLFEAKHLKHSYITVANTGSRPIVVPANGYFRFDTWVWQFYSLAPMG 526  
QY 540 TNGRRRIQ 548  
Db 527 TNGRRRRVQ 535  
RESULT 5  
AAB49706  
ID AAB49706 standard; protein; 540 AA.  
XX AAB49706;  
AC AAB49706;  
DT 04-APR-2001 (first entry)  
DE Small round structured virus protein SEQ ID 7.  
KW Small round structured virus; SRSV; food poisoning.  
OS Small round structured virus.  
PN WO200079280-A1.  
PD 28-DEC-2000.  
PF 22-JUN-2000; 2000WO-JP004095.  
PR 22-JUN-1999; 99JP-00175928.  
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX WPI; 2001-080848/09.  
DR N-PSDB; AAF29147.  
XX Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX Claim 1; Page 54-57; 84pp; Japanese.  
PS This invention relates to a kit for the detection and typing of small  
XX round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX  
SQ Sequence 540 AA;  
Query Match 69.7%; Score 2018; DB 4; Length 540;  
Best Local Similarity 68.3%; Pred. No. 1.4e-169;  
Matches 377; Conservative 71; Mismatches 88; Indels 16; Gaps 4;  
QY 1 MKVASNDAPSNQAGLVPESNNEAMALPEVVGASLAAPVTGQTNIIDPWIRTNFVQAP 60

Db 1 MKVASNDAPSNQAGLVPESNNEAMALPEVVGASLAAPVTGQTNIIDPWIRTNFVQAP 60  
QY 61 GGEFTVSPRNPGEVLLNLELGPINPYLAHLARMYNGYAGGFVQVVLGAGNAFTAGKII 120  
Db 61 NGEFTVSPRNPGEVLLNLELGPINPYLAHLARMYNGYAGGFVQVVLGAGNAFTAGKII 120  
QY 121 FRAIIPNPPIDNLSAAQITMCPHVIVDVROLEPNLPMPPVRRNPPFHYNQSGDSRLRLIA 180  
Db 121 FRAVPPYFFVENLSPQITMFPVHIIIDVRLTLEPVLPMPPVRRNPPFHYNQSGDSRLRLIA 180  
QY 181 MLYTPLRANNSGDDVFTVSCRVLTRPSPDFSNELVPTVESKTPFTLPILTISEMSNS 240  
Db 181 MLYTPLRANNSGDDVFTVSCRVLTRPSPDFSNELVPTVESKTPFTLPILTISEMSNS 240  
QY 241 RFPVPTESLHTSPTENIVVQCONGRVTLTGELMGTQTLPSQICAFRGLVLRSTRSRADQ 300  
Db 241 RFPPLSIDEMVTSFNESIVVQPNQGRVTLTGELMGTQTLPSQICAFRGLVLRSTRSRADQ 300  
QY 301 ADTATPRLFNYYVHVLQNLNGTPYDPAEDIPGPIGTDPDRGKVFVGASQRNLDSTTRAH 360  
Db 298 -----NHWNQVNTNINGTPDPTGDPAPLGTDFSGKLFVLSQRDHDAACKSH 348  
QY 358 RAHEAKVDTTAGRTFPTKLGSLLEIST-DSDDFDQNPQTKFTPTVGVIG-VDNBAEFQOWSLPDY 416  
Db 349 RAHDVAVATYSDKYTPKLGVLQIGTWTNDVENVQTKFTPIGLNEVANGHRFEQWTLPRY 408  
QY 417 SQQFTHNMNLAPAVAPNPFGEQQLLFFRSQSSGRSNGVLDCLVPOEWVQHEYESAP 476  
Db 409 SGALTLMNLAPAVAPLFGCERLLFFRSYVPLKGGFNPDAIDCSVPQEWVQHEYESAP 468  
QY 477 QTQVALVRYVNPDTGKVLFEAKHLKGFMTIANNGDSPTVPPNGYFRFESWVAPFYTLA 536  
Db 469 LGDVALVRYVNPDTGKVLFEAKHLKGGFLTVSSTIGTPVVPVVPANGYFRFDSWVQFYSLA 528  
QY 537 PMGTGNGRRRIQ 548  
Db 529 PMGTGNGRRRVQ 540  
RESULT 6  
AAB49704  
ID AAB49704 standard; protein; 539 AA.  
XX AAB49704;  
DT 04-APR-2001 (first entry)  
DE Small round structured virus protein SEQ ID 5.  
KW Small round structured virus; SRSV; food poisoning.  
OS Small round structured virus.  
PN WO200079280-A1.  
PD 28-DEC-2000.  
PF 22-JUN-2000; 2000WO-JP004095.  
PR 22-JUN-1999; 99JP-00175928.  
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX WPI; 2001-080848/09.  
DR N-PSDB; AAF29145.  
XX Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX



PS Claim 1; Page 50-52; 84pp; Japanese.

XX This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks

XX Sequence 539 AA;

DR Query Match 67.2%; Score 1947.5; DB 4; Length 539;  
 DR Best Local Similarity 66.2%; Pred. No. 2.5e-163;  
 XX Matches 365; Conservative 73; Mismatches 94; Indels 19; Gaps 4;

Qy 1 MKWASNDAPSNDGAGLVEPEINNEAWALDPVAGAAIAPLTGQONIIDPWIMNNFVQAP 60  
 Db 1 MKWASNDAPSNDGAGLVEPEINNEAWALDPVAGAAIAPLTGQONIIDPWIMNNFVQAP 60

Qy 61 GGEFTVSPRNSPGEVLINLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAFTAGKII 120  
 Db 61 GGEFTVSPRNSPGEVLINLELGPENPYLAHLARMYNGYAGGFEVQVVLGNAFTAGKII 120

Qy 121 FAAPNPFIDNLISAAQITWCPHVIDVRLQLEPVLNLPMDVVRNPFHYNQSGDSRLRLIA 180  
 Db 121 FAAPNPFIDNLISAAQITWCPHVIDVRLQLEPVLNLPMDVVRNPFHYNQSGDSRLRLIA 180

Qy 181 MLTYPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTPKPTLPILTISEMSNS 240  
 Db 181 MLTYPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTPKPTLPILTISEMSNS 240

Qy 241 RFPVPIESLHTSPTEINIVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTSTRSRASDQ 300  
 Db 241 RFPVPIESLHTSPTEINIVQCONGRVTLTGELMGTTQLLPSQICAFRGVLTSTRSRASDQ 300

Qy 301 ADTATPRLFNYYHVVQDLNLTGTPYDPAEDIPGLGTDPFRGKVFVGASQRNL-DSTTRA 359  
 Db 301 ADTATPRLFNYYHVVQDLNLTGTPYDPAEDIPGLGTDPFRGKVFVGASQRNL-DSTTRA 359

Qy 467 AQSDVALLRVNPDTRGLVFECKLHKSQVYVTAHTGPHDLVIPPNGYFRFDSVWNQFVTL 526  
 Db 467 AQSDVALLRVNPDTRGLVFECKLHKSQVYVTAHTGPHDLVIPPNGYFRFDSVWNQFVTL 526

Qy 536 APMTGTGRRR 546  
 Db 536 APMTGTGRRR 546

Qy 527 APMGNGAGRRR 537  
 Db 527 APMGNGAGRRR 537

RESULT 7  
 AAB49708  
 ID AAB49708 standard; protein; 542 AA.  
 XX AAB49708;  
 XX AAB49708;  
 XX 04-APR-2001 (first entry)  
 XX Small round structured virus protein SEQ ID 9.  
 XX Small round structured virus; SRSV; food poisoning.  
 XX Small round structured virus.  
 XX WO200079280-A1.  
 XX

RESIST 7

AAB49708

ID AAB

XX

AC AAB

3

5

DI  
04-  
yy

DE SMA

2000

KW Sma

XX

**OS** **Sma**

33

PN  
WQZ  
yy

W

Db 300 -----YRMDIEITNDGTPIDTDPGPIGSPDFQILFGVASQKNEQNP 348  
 QY 357 TRAHEAKVDTTAGRTFKLGLSLEISDSDDFDQ-NQPTKFTFVG-----IGVDNEAEFQ 409  
 Db 349 TRAHEAIINTGGDHLCPQISSEIYLTSPNLRCTNPNQPLQOSGLRGTILIRSDN-----403  
 QY 410 QMSLPDYSQGFTHNNLAPAVAPNPPGQLLFFRSQSPSSGSRNGVLDCLVPQEWQHF 469  
 Db 404 -----GHCHDMVGTSTPTTPWQWRRCSSRGSSGHRYPPVPMNRVTMIVLS 454  
 QY 470 YOBESAPAQTV-----ALVRYVNPDTGKVLFAKLHKLGMFTIANGDSPIVPPNG 521  
 Db 455 HKSGFSTIRKLPQLNLRWPLIRFINPDTGRVLFEARLHKQGFITVAHTGDNPIVMPNG 514  
 QY 522 YRFESWVNPFFYTLAPMGTGNGRRRIQ 548  
 Db 515 YRFEAWNQFYSLAPVGTGKRRRVQ 541

RESULT 9  
 AAR50972 standard; protein; 530 AA.  
 XX AAR50972;  
 AC AAR50972;  
 XX 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 05-OCT-1994 (first entry)  
 XX Norwalk virus strain 8FIIa protein (encoded by ORF2).  
 XX Norwalk virus; pathogen; acute gastroenteritis; food poisoning;  
 KW seafood contamination; diagnostic assay; calicivirus; small round virus.  
 XX Norwalk virus; (strain 8FIIa).  
 XX WO9405700-A2.  
 XX 17-MAR-1994.  
 XX 07-SEP-1993; 93WO-US008447.  
 XX 07-SEP-1992; 92US-00941365.  
 XX (BAYU ) BAYLOR COLLEGE MEDICINE.  
 XX Matson DO, Estes MK, Jiang X, Graham DY;  
 XX WPI; 1994-101125/12.  
 DR N-PSDB; AAQ56826.  
 XX DNA from Norwalk and related viruses - used for preparing prods. for use  
 PT in diagnostic assays, detection and vaccines for Norwalk and related  
 PT viruses.  
 XX Claim 14; Page 68-70; 156pp; English.  
 XX The Norwalk virus was isolated from stool samples from adult volunteers  
 CC infected with safety tested Norwalk virus strain 8FIIa. The coding  
 CC sequence is useful for the design of probes for use in diagnostic assays  
 CC for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct  
 CC FN field.) (Updated on 16-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 530 AA;

Query Match 41.8%; Score 1211; DB 2; Length 530;  
 Best Local Similarity 45.4%; Pred. No. 4.1e-98;  
 Matches 260; Conservative 74; Mismatches 161; Indels 78; Gaps 11;

QY 1 MKVASNDAAAPSDGAG---LVPEIN-NEAMALDPVAGAAIAAPLTGQNIIDPIMNNF 56

Db 1 MMASKDQATSSVDGASGAGQLVPEVNASDPLANDPVAGSSTAVATAGVNPDPIMNNF 60

RESULT 8  
 ID AAB49710 standard; protein; 541 AA.  
 XX AAB49710;  
 AC AAB49710;  
 DT 04-APR-2001 (first entry)  
 XX Small round structured virus protein SEQ ID 11.  
 DE Small round structured virus; SRSV; food poisoning.  
 XX Small round structured virus.  
 KW Small round structured virus.  
 XX Small round structured virus.  
 OS Small round structured virus.  
 XX WO200079280-A1.  
 PN 28-DEC-2000.  
 PD 22-JUN-2000; 2000WO-JP004095.  
 XX 22-JUN-1999; 99JP-00175928.  
 PR (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 PA (DENK-) DENKA SEIKEN KK.  
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 FI WPI; 2001-080848/09.  
 DR N-PSDB; AAF29151.  
 XX Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.  
 XX Claim 1; Page 64-66; 84pp; Japanese.  
 XX This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX  
 SQ Sequence 541 AA;  
 Query Match 58.0%; Score 1678.5; DB 4; Length 541;  
 Best Local Similarity 58.9%; Pred. No. 1.7e-139;  
 Matches 334; Conservative 67; Mismatches 121; Indels 45; Gaps 7;

QY 1 MKVASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPIMNNFVQAP 60  
 Db 1 MKVASNDAAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNIIDPIMNNFVQAP 60  
 QY 61 GGEFTVSPRNSPCEVLLNLELPEINPYLAHARMVNGVAGFEVQVVLGNAFTAGKII 120  
 Db 61 AGEFTVSPRNSPGEIILDELGPDLPYLAHARMVNGHAGGEVQIVLAGNAFTAGKII 120  
 QY 121 FRAIPNFPIDNLAAQITMCQHVIVDRQLEFVNLPMFDPVARNFPHYNNQSGDSRLRLIA 180  
 Db 121 FRAIPPGFFENYLSQITMCQHVIVDRQLEFVNLPMFDPVARNFPHYNNQSGDRLRLIA 180  
 QY 181 MLTYPLRANNSGDDVTVSCRVLTRSPDPSFNFLVPTVESKTKPFTLPILTISEMNS 240  
 Db 181 MLTYPLRANNSGDDVTVSCRVLTRSPDPSFNFLVPTVESKTKPFTLPILTISEMNS 240  
 QY 241 RFPVPIESLHTSPENIVWQCONGRVTLDGELMGTQLLPSQICAFRG-VLTRFSTRASD 299  
 Db 241 RFPVPVDMVTARNENQVQPNQGRVTLDGELMGTQLLPSQICAFRG-VLTRFSTRASD 299  
 QY 300 QADTATPRLFNYYWHVQDLNGLNGTPVDPDAEDIPGLGTDPDFRGKVFVGASQRLNDS---T 356



OS Small round structured virus.  
XX WO200079280-A1.  
XX 28-DEC-2000.  
XX 22-JUN-2000; 2000WO-JP004095.  
XX 22-JUN-1999; 99JP-00175928.  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX (DENK-) DENKA SEIKEN KK.  
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX WPI; 2001-080848/09.  
XX N-PSDB; AAF29142.  
XX Kit for the detection and typing of small round-structured virus (SRSV)  
XX strains for investigation of food poisoning outbreaks, contains  
XX antibodies.  
XX Claim 1; Page 42-45; 84pp; Japanese.  
XX This invention relates to a kit for the detection and typing of small  
XX round structured virus (SRSV) strains. The kit contains antibodies  
XX directed against peptides represented in sequences AAB49700 - AAB49710,  
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
XX used for detecting and typing strains of SRSV in order to prevent the  
XX spread of infection and to examine the epidemiology of outbreaks  
XX  
XX Sequence 530 AA;  
Query Match 41.7%; Score 1209; DB 4; Length 530;  
Best Local Similarity 45.4%; Pred. No. 6.1e-98;  
Matches 260; Conservative 72; Mismatches 163; Indels 78; Gaps 11;  
QY 1 MKMASNDAPSNDAAG---LVPEIN-NEAMALDPVAGAAIAAPLTGQNIIDPWIMNPF 56  
DB 1 MMASKDATSVDGASGAGQLVEVNASDPLAMDVAGSSTAVATAGQVNPIDPWIMNPF 60  
QY 57 VOAPGGFTVSPRNSPGEVLLNLELGPENIPYLAHARMYNGYAGGFEVQVVLGNAFTA 116  
DB 61 VOAPQGEFTISPNTPGGVLFDLGLGPHLAPFLHLSQMYNGVGNVRVIMLAGNAFTA 120  
QY 117 GKIIFAAIPNFPIDNLSAAQITMCPHVIVDVROLEPNLPMPPVNRNFFHYNQSDSRL 176  
DB 121 GKIIIVSCIPFGGSHNLTIQAQLFPHVIADVRTLDPIEVLVDVNRVLFHNDRNQOTM 180  
QY 177 RLIAMLYTPLRANSGDDVFTVSCRVLTRSPDFSNFLVPPPTVESKTKPTLPILTISE 236  
DB 181 RLVCMLYTPLRGTGGTGSFVYAGRVMTCPSPDFNLFVLVPPPTVEQKTRFTLPLNPLSS 240  
QY 237 MNSRFPVPIESLHTSPENTVVOCONGRVTLDCGLMGTQTLPLSQICAFRGVLTSTR 296  
DB 241 LSNRAPLPISGMISPDNVQSVQFQNGRCRTLGLVGTTPVLSLHVAKIRG-----TSN 295  
QY 297 ASDQADTATPLFNYMHVOLDNLNGTPYDPAEDIPGLGTPDPFG-----KVFGVAS 349  
DB 296 GT-----VINLTDLGTPFHPFEG-PAPIGFPLGCGDWHINMTQFGHSS 339  
QY 350 QBNLDSSTRAEAKVDTTAGFTPKLGSLEISTDSDDFDQNKPTKFTPVGIGVDN----- 404  
DB 340 QTYD-----VDTTDPFVPHLGSIQAN-----GIGSGNYIGVL 373  
QY 405 -----EAEFQOWSLPDYSGQFTNNMLNAPAVAPNPFGEQLLFFRSQPSGGRSN 454  
DB 374 SNVSPSPHSGQVDLMKIPNYGSSI TEATHLAPSVYPPGFEVLVFPFMSKIPGFGAYS- 432  
QY 455 GVLDCILVQEWVQHYQESAPAQTOVALRVYVNPDTGKVLFEAKLHKLGFMTIANNGDS- 513  
DB 433 --LPCLLPQEIYSHLASEQAFTVGEAALLHYVDPDTGTGLGFEKAYPDGFLTCVPGNASS 490

QY 514 -PITVPPNGYFRFESWVNPFFYTLAPMGTCNGRR 545  
DB 491 GPQQLPFGVFFVSWSRFYQLKPKVGTASSAR 523  
RESULT 12  
AAB49700  
ID AAB49700 standard; protein; 545 AA.  
XX AAB49700;  
XX 04-APR-2001 (first entry)  
XX Small round structured virus protein SEQ ID 1.  
XX Small round structured virus; SRSV; food poisoning.  
XX Small round structured virus.  
XX WO200079280-A1.  
XX 28-DEC-2000.  
XX 22-JUN-2000; 2000WO-JP004095.  
XX 22-JUN-1999; 99JP-00175928.  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX (DENK-) DENKA SEIKEN KK.  
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX WPI; 2001-080848/09.  
XX N-PSDB; AAF29141.  
XX Kit for the detection and typing of small round-structured virus (SRSV)  
XX strains for investigation of food poisoning outbreaks, contains  
XX antibodies.  
XX Claim 1; Page 40-42; 84pp; Japanese.  
XX This invention relates to a kit for the detection and typing of small  
XX round structured virus (SRSV) strains. The kit contains antibodies  
XX directed against peptides represented in sequences AAB49700 - AAB49710,  
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
XX used for detecting and typing strains of SRSV in order to prevent the  
XX spread of infection and to examine the epidemiology of outbreaks  
XX  
XX Sequence 545 AA;  
Query Match 40.7%; Score 1178.5; DB 4; Length 545;  
Best Local Similarity 45.7%; Pred. No. 3.2e-95;  
Matches 258; Conservative 82; Mismatches 183; Indels 41; Gaps 14;  
QY 1 MKMASNDAPSNDAAG---LVPEIN-NEAMALDPVAGAAIAAPLTGQNIIDPWIMNPF 56  
DB 1 MMASKDAPTNDGTSGAGQLVPEANTAEPISTMEPVAGAAATAATAGQVNPIDPWIMNY 60  
QY 57 VOAPGGFTVSPRNSPGEVLLNLELGPENIPYLAHARMYNGYAGGFEVQVVLGNAFTA 116  
DB 61 VOAPQGEFTISPNTPGDILFDLQGLPHLAPFLSHLAQMYNGVGNMKVKVLLAGNAFTA 120  
QY 117 GKIIFAAIPNFPIDNLSAAQITMCPHVIVDVROLEPNLPMPPVNRNFFHYNQSDSRL 176  
DB 121 GKIIISCIPEGFAAQNISIAQATMFPFHVIADVRVLEPIEVLVDVNRVLFHNDRNAPT-M 179  
QY 177 RLIAMLYTPLRANSGDDVFTVSCRVLTRSPDFSNFLVPPPTVESKTKPTLPILTISE 234  
DB 180 RLVCMLYTPLRASGSSGTDPPVIAGRVLTCSPPDFSLFLVPPNVEQTKFSPVNLPL 239  
QY 235 SEMSNRFPVPIESLHTSPENTVVOCONGRVTLDCGLMGTQTLPLSQICAFRGVLTSTR 294

Db 240 NTLSNRVPSLIKSMVSRDHQVQFQNGRVTLDGQLQGTTPTASQLCKIRGSVFHAN 299  
 QY 295 SRASQADATATPRLFNYYWHVQDNLNGTTPYDPAEDIPGLGTPDFRGKV-----FGVAS 349  
 Db 300 GG-----NGY---NLTELDGSPYHAFES-PAPIGFPDL-GECDWHMEASPTT 341  
 QY 350 QRLNSTTTRAHAKVDTTAGRTPKLGSLEISTDS-DFDQNPYKF---TPVGIGVDNE 405  
 Db 342 QNTGDIKQINVKQESA---FAPHLGTIQADGLSDVSNTMIKLGWSPVSDG--HR 396  
 QY 406 AEFQWMLSDYSGQTHNNMLAPAVAPNPPGQOLLFFERSQLPSSGSGRNGVLDCLVROEW 465  
 Db 397 GVDVFWVPRYGSTITEAQAAPPYPPGFGGAIVFFMSDFPIAGHTNGLSVPCITPOEF 456  
 QY 466 VOHFYQESAPQTOVALVRYNPNFTGKLVFEAKHLKGLFMTIANN--GDSPTITPPNGYF 523  
 Db 457 VTHFVNEQAPTRGEAALLHYLDPDTHRLNGEFLYPEGFMTCVPSNGSGTGPTQLPINGVF 516  
 QY 524 RPSWNPFTYLA PMGTGNRRRI 547  
 Db 517 VFVSWVSRYQLKPVGTAGPACRL 540

## RESULT 13

ID AARS7091  
 XX AARS7091 standard; protein; 530 AA.

AC AARS7091;

DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 05-OCT-1994 (first entry)

XX Small round virus SRSV/KY/89 capsid protein.

XX pathogen; acute gastroenteritis; food poisoning; seafood contamination;

KW diagnostic assay; human calcivirus; small round virus; SRSV; KY89;

KW Norwalk virus; capsid protein.

XX Small round structured virus.

XX WO9405700-A2.

XX 17-MAR-1994.

PF 07-SEP-1993; 93WO-US008447.

XX 07-SEP-1992; 92US-00941365.

XX (BAYU ) BAYLOR COLLEGE MEDICINE.

XX Matson DO, Estes MK, Jiang X, Graham DY;

XX WPI; 1994-101125/12.

DR N-PSDB; AAQ56832.

PT DNA from Norwalk and related viruses - used for preparing prods. for use  
 PT in diagnostic assays, detection and vaccines for Norwalk and related  
 PT viruses.

PS Example 7; Fig 13a; 156pp; English.

CC The known sequence for Norwalk virus was used to obtain the sequence of  
 CC other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool  
 CC from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide  
 CC cDNA sequence includes part of the polymerase region and the capsid  
 CC region of the genome; the deduced amino acid sequences are AARS7092 and  
 CC AARS7091, respectively. Expression of fragments and derivs. of Norwalk-  
 CC related viruses permits development of diagnostic assays to detect  
 CC antibodies, antigens, viral genetic material or antivirals. (Updated on  
 CC 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS  
 CC field.)

XX SQ Sequence 530 AA;  
 Query Match 40.2%; Score 1164; DB 2; Length 530;  
 Best Local Similarity 44.4%; Pred. No. 5, 9e-94;  
 Matches 255; Conservative 74; Mismatches 165; Indels 80; Gaps 13;  
 QY 1 MKMASNDAAPSNDGAG---LVPEIN-NEAMALDPVAGAAIAAPJTQONIIDPIMMNP 56  
 Db 1 MMASKDATSSVDCASASVQLVPEVNASDPLAMDPAVAGSTAVATAGVNPIDPDIINF 60  
 QY 57 VOAPGGEFTVSPRSPGVLNLEIGPEINPVLAAHARMYNGYAGGFEVQVVLGNAFTA 116  
 Db 61 VOAPGGEFTIISPNNTPGDVLFDLSLGPHLNPFLLHLSQMYNGWGNMVRIMLAGNAFTA 120  
 QY 117 GKIIIPAAIPPPFPIDNLSAAQITMCCHPVIIVDVRQLEPNLMPDPVRRNPFHYNOGSDRL 176  
 Db 121 GKIIIVSCIPPGSGSQTIAQATLFFHVIADVRLDPIEVLDPVRLVFNHNDNRNQTM 180  
 QY 177 RLIAMLYTPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPTVESKTKPFTLPILTISE 236  
 Db 181 RLVCMLYTPLSTGGTGDGDFVAVGRVMTCPSPDFNFLVPPTVBQKTRPFTLPNPLSS 240  
 QY 237 MNSRPPVPVPIESLHTSPTEINIVQCGRVTLDGELMGTTLQLLPSQICAPRGVLTSTRSP 296  
 Db 241 LNSRAPLPISGMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRG-----TSN 295  
 QY 297 ASDQADTATPRLFNYYWHVQDNLNGTTPYDPAEDIPGLGTPDFRG-----KVFGVAS 349  
 Db 296 GT-----VINLTEDGTFFHPEG-PAPIGFPDLGGCDWHNMTOFGHSS 339  
 QY 350 QRLNSTTTRAHAKVDTTAGRTPKLGSLEISTDSDFDQNPYKFPTVIGVDN-----404  
 Db 340 QTQYD-----VDTTPTDSVPHLGSIQAN-----GIGSGNYIGVL 373  
 QY 405 -----EAEFOQWSLPDYSQFTNMMLAPAV-APNPFGEQLLFFRSOLPSSGGRS 453  
 Db 374 SWSPPSPHSPSGQVDLMKIPNYGSSITBATHLAPSVSPGF-GEVLVFFMSKIPGPGDS 432  
 QY 454 NGVLDCLVPOQSVQVHFYQESAPQTOVALVRYNPNFTGKLVFEAKHLKGLFMTIANN 513  
 Db 433 ---LPCLLPQGYIISHLASEQAPTVEGGLLHYVDPTDNRNLGFEKAYPDGFLTCVPNGAS 489  
 QY 514 --PITVPENGYFRFESWYNPFYTLAPMTGNRR 545  
 Db 490 SGPOQLPINGVFVFWVSWSRYQLKPVGTASTAR 523  
 RESULT 14  
 AAB49703  
 ID AAB49703 standard; protein; 544 AA.  
 XX AC AAB49703;  
 XX 04-APR-2001 (first entry)  
 XX Small round structured virus protein SEQ ID 4.  
 DE Small round structured virus; SRSV; food poisoning.  
 KW Small round structured virus.  
 OS Small round structured virus.  
 PN WO200079280-A1.  
 XX PD 28-DEC-2000.  
 XX PF 22-JUN-2000; 2000WO-JP004095.  
 XX PR 22-JUN-1999; 99JP-00175928.  
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 PA (DENK-) DENKA SEIKEN KK.  
 XX

Wed Jun 2 09:13:30 2004

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 XX WPI; 2001-080848/09.  
 DR N-PSDB; AAF29144.  
 XX  
 DR Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.  
 XX  
 PS Claim 1; Page 47-49; 84pp; Japanese.  
 XX  
 CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX  
 SQ Sequence 544 AA;  
 Query Match 39.5%; Score 1143.5; DB 4; Length 544;  
 Best Local Similarity 44.3%; Pred. No. 4e-92;  
 Matches 252; Conservative 86; Mismatches 183; Indels 45; Gaps 15;  
 QY 1 MKVASNDAPSDGAG---LVPEINN-EMALDPVAGAAIAAPLTGQONIIDPWIMNF 56  
 Db 1 MMASKDATSADGATGAGQLVPEVNTADPIIDPVAGSSTALATAGQVNLIDPWIMNF 60  
 QY 57 VQAPGGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGGEVQVVLGNAFTA 116  
 Db 61 VQAPGGEFTISPNNTPGDILFDLQGLPHLNFPLSHLSQMYNGWGNRVRVVLGNAFTA 120  
 QY 117 GKIIFAAIPNFPIDNLSAQITMCPHVIDVRQLEPNLMPDVRNFFHYNQGSRL 176  
 Db 121 GKVIICCPGFGQRTSLIAQATLPHVIADVRLEDPVEPLEVRNVLH-NNDTQPTM 179  
 QY 177 RLIAMLYTPLRAN--NSGDDVFTVSCRVLTSPDPSFNFVLPVPTVESKTPFLPILTI 234  
 Db 180 RLLCMLYTPLTGCGASGTDTSFVAGRVLTCPGDFNLFVLPVPTVEQKTRPFTVFNIP 239  
 QY 235 SEMSNRPVPIESLHTSPENIVVQCONGRVTLDBGLMGTQALLPSQICAFRGVLTST 294  
 Db 240 KYLSNSRIPNPIEGNSLSPDQTNVQFNGRCTIDGQPLGTPVSVLSQCKFRGRI---- 295  
 QY 295 SRASDAQDTATPRLFNYYWHVOLNLTNGTPYDPAEDIPGLCTPDPRGKVFVASQRNLD 354  
 Db 296 -----TSQORVLN-----LTELDSGF-MAFAAPAPAGFPDLGSCDWHIEMSKIPN 340  
 QY 355 STTRAHEA---KYDTTAGRFTPKLSL-----EISTDSDDFDQNPQTKFTPVGIGVDNEAE 407  
 Db 341 SSTQNNPIVTSVKNSQOQVPHLSLITLDENSVSGDYIGTIQWTS-PPSDSG---GAN 396  
 QY 408 FQWLSLPDYSGQFTHNMMLAPAVNPFCEQLLFFRSQLP---SSGRNSGVLDCLVPOE 464  
 Db 397 TNFWKIPDYGSSLAESQAPAVYPGGEVNEIVYFEMASIPGNOSG--SPNLVFCLLPQE 454  
 QY 465 WYOHFQESAPQATQVALRVYNPDGKVLPAKHLKLGFWMTIANGDS--PITVPPNGY 522  
 Db 455 YITHFTISEQAPIQGEAALHYDPTNRNLGFEKLYPGGYLTVCVPSNSSSTGPOQLPDGV 514  
 QY 523 PFESWVNPFTLAPMGT--GNRRRI 547  
 Db 515 FVFASWVSRYQLKPVGTAGPARGRL 540  
 RESULT 15  
 AAB49702  
 ID AAB49702 standard; protein; 546 AA.  
 XX  
 AC AAB49702;  
 XX  
 DT 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 3.  
 DE  
 XX Small round structured virus; SRSV; food poisoning.  
 KW  
 XX Small round structured virus.  
 OS  
 XX WO200079280-A1.  
 PN  
 XX 28-DEC-2000.  
 PD  
 XX 22-JUN-2000; 2000WO-JP004095.  
 XX  
 PF 22-JUN-1999; 99JP-00175928.  
 XX  
 PR (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 XX  
 PA (DENK-) DENKA SEIKEN KK.  
 XX  
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 XX  
 XX WPI; 2001-080848/09.  
 DR N-PSDB; AAF29143.  
 DR  
 XX Kit for the detection and typing of small round-structured virus (SRSV)  
 XX strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.  
 PT  
 XX Claim 1; Page 45-47; 84pp; Japanese.  
 PS  
 XX This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 CC  
 XX Sequence 546 AA;  
 SQ  
 Query Match 38.8%; Score 1124.5; DB 4; Length 546;  
 Best Local Similarity 43.0%; Pred. No. 1.9e-90;  
 Matches 250; Conservative 80; Mismatches 182; Indels 69; Gaps 17;  
 QY 1 MKMASNDAPSDGAG---LVPEINN-EMALDPVAGAAIAAPLTGQONIIDPWIMNF 56  
 Db 1 MMASKDAPQADGAGAGQVLPVNTADPLPMEPVAGPTTAVATAGVNMIDPWIMNF 60  
 QY 57 VQAPGGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAGGEVQVVLGNAFTA 116  
 Db 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLNFPLSHLSQMYNGWGNRVRVILLAGNAFSA 120  
 QY 117 GKIIFAAIPNFPIDNLSAQITMCPHVIDVRQLEPNLMPDVRNFFHYNQGSRL 176  
 Db 121 GKIIIVCCVPPGFTSSSLTAQATLPHVIADVRLEPIEMPLEDVRNVLVHTNDNQPT-M 179  
 QY 177 RLIAMLYTPLR--ANNSGDDVFTVSCRVLTSPDPSFNFVLPVPTVESKTPFLPILTI 234  
 Db 180 RLVCMLYTPLTGCGSGNSDSFVAGRVLTAPSDSFSLFVLPPTIEQKTRFTVPTPL 239  
 QY 235 SEMSNRFPVPIESLHTSPENIVVQCONGRVTLDBGLMGTQALLPSQICAFRGVLTST 294  
 Db 240 QTLNSRFPFLIQMILSPDASQVVFQNGRCLIDGQLLGTTPATSGQLFVRVGKINQGA 299  
 QY 295 SRASDAQDTATPRLFNYYWHVOLNLTNGTPYDPAEDIPGLCTPDPRGK----VFGVASQ 350  
 Db 300 -----RTLN-----LTEVDGKFP-WAFDSPAIVGFPDF-GKCDWHMRISKTP 339  
 QY 351 RNLDSSTTRAHEAKVDTTAGRFTPKLSLEISTDSDFDQ--NQPT-----KFT 396  
 Db 340 NNTSSGDPMSVSQVQTNVQGFVPHLSIQ-----FDEVFNHPTGDTYTGTEWISQPT 392  
 QY 397 PVGIGVDNEAEFQOWSLPDYSGQFTHNMMLAPAVNPFCEQLLFFRSQPSGGRS--N 454

Db 393 PPG-----TDINLWEIPDYSSLSQAANLAPPVFPFGFGEALVYFVSAPFGPNRSPN 446  
QY 455 GVLDCLPQEWYQHFYQESAPAQVALVRYVNPDTGKVLFEAKLHKLGFMTIANN--GD 512  
Db 447 DV-PCLLPQEXYTHFVSEQPTMGDAALLHYVDPDTNRNLGEFKLYPGGYLTCVENGVA 505  
QY 513 SPITVPPNGYFRFESWVNPFFYTLAPMGTGN-----GRRRI 547  
Db 506 GFQQLPLNGVFLFVSWSRFYQLKEVGTASTARSLGVRI 546

Search completed: June 1, 2004, 13:45:54  
Job time : 49.0765 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 13.0564 Seconds  
(without alignments)  
2166.837 Million cell updates/sec

Title: US-09-926-799-6  
Perfect score: 2896  
Sequence: 1 MKMASNDAPSDGAGLVLP.....VNPFYTLAPMGNGRRRIQ 548

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1211	41.8	530	4	US-08-486-049-3
2	285.5	9.9	622	4	US-09-590-020-6
3	281	9.7	623	4	US-09-590-020-4
4	279.5	9.7	626	4	US-09-590-020-7
5	279.5	9.7	668	4	US-09-617-594A-4
6	276	9.5	623	4	US-09-590-020-2
7	262	9.0	669	4	US-09-617-594A-2
8	128.5	4.4	2137	4	US-09-134-001C-4463
9	121	4.2	776	3	US-08-961-083-86
10	121	4.2	776	4	US-09-536-784-86
11	119	4.1	2318	3	US-09-091-219-24
12	119	4.1	2318	4	US-09-660-541-24
13	117	4.0	2206	1	US-07-852-260-2
14	117	4.0	2206	2	US-08-461-503-2
15	117	4.0	2206	3	US-08-465-250-2
16	114.5	4.0	3892	4	US-09-328-352-5503
17	108.5	3.7	883	4	US-09-489-039A-13542
18	107.5	3.7	1043	3	US-08-928-361B-30
19	107.5	3.7	1721	3	US-08-700-651-5
20	107.5	3.7	1721	3	US-08-928-361B-6
21	107.5	3.7	1721	4	US-09-588-995A-6
22	104	3.6	1257	2	US-08-750-152A-2
23	102	3.5	540	4	US-08-956-171E-5225
24	102	3.5	540	4	US-08-956-171E-5242
25	102	3.5	1560	4	US-09-264-512B-2
26	101	3.5	1447	3	US-09-041-886-25
27	101	3.5	1447	5	PCT-US94-05277-2

28	101	3.5	2232	3	US-09-091-219-25	Sequence 25, Appl
29	101	3.5	2232	4	US-09-660-541-25	Sequence 25, Appl
30	101	3.5	2247	3	US-09-091-219-2	Sequence 2, Appl
31	101	3.5	2247	4	US-09-660-541-2	Sequence 2, Appl
32	100.5	3.5	3788	4	US-09-336-447A-76	Sequence 76, Appl
33	99.5	3.4	968	4	US-09-228-986-76	Sequence 76, Appl
34	97	3.3	1042	3	US-08-928-361B-11	Sequence 11, Appl
35	97	3.3	1042	4	US-09-588-995A-11	Sequence 11, Appl
36	97	3.3	1837	3	US-08-928-361B-5	Sequence 5, Appl
37	97	3.3	1837	4	US-09-588-995A-5	Sequence 5, Appl
38	96.5	3.3	1056	4	US-09-079-030-217	Sequence 217, Appl
39	96.5	3.3	4536	4	US-09-180-432B-27	Sequence 27, Appl
40	96.5	3.3	4536	4	US-09-079-030-1	Sequence 1, Appl
41	95.5	3.3	401	4	US-08-858-207A-344	Sequence 344, Appl
42	95.5	3.3	447	4	US-09-627-376-10	Sequence 10, Appl
43	95	3.3	501	2	US-08-448-603A-2	Sequence 2, Appl
44	95	3.3	501	3	US-09-134-075-2	Sequence 2, Appl
45	95	3.3	501	4	US-09-492-739-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-08-486-049-3  
; Sequence 3, Application US/08486049  
; Patent No. 6572862  
; GENERAL INFORMATION:  
; APPLICANT: Bates, Mary K  
; APPLICANT: Jiang, Xi  
; APPLICANT: Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSE: Fulbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE:  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-486-049-3

Query Match 41.8%; Score 1211; DB 4; Length 530;  
Best Local Similarity 45.4%; Pred. NO. 4.8e-119;  
Matches 260; Conservative 74; Mismatches 161; Indels 78; Gaps 11;  
Qy 1 MKMASNDAPSDGAG---LVPEIN-NEAMALDPVACAAIAAPLTGQNIIDPHMNF 56  
DB 1 MMASKDATSSVDGASGAGQLVPEYNASDPLAMDPAVAGSTAVATAGQVNPIDPWIINF 60

```

186 LRANNSGDDVETVSC--RVLTRPSPDFSNFLVPTVBESKTKPFTLPILTISEMSNSRFP 243
250 L-INPYANTNMSGGIVVEYKPGDFKHLKPPG-----SMLTHGSVPDILP 298
244 ----VPIESLHTSPTEINIVVQNGRVTLDGELMGTTLQSPQICAFRGVLTSTRSRASD 299
299 KSSSLWIGNRHSWDTDFIIR-----PFVQANRHFDFNQE 334
300 QADTATPRLFNYYWHVQDLNLTGTPVDAED--IPG-PLGTDPDR-----GKVFVGAS 349
335 TAGWSTPRFRPITITVSESNMKSGLIGVATDVIIVGIPDGNPDTTIPEQLTPAGIYSITA 394
350 QRNLDSGTR----AHEAKVDTTAGRTPKLGSLEISTDSDDDQNOPTKFTPVGIGVDNE 405
395 SNGTDTITTAAGYDAAETIVNTTNFKSMYICGSLQRAWDDKKISN---TAFITTAVRKGN 451
406 AE 407
452 IE 453

RESULT 3
US-09-590-020-4
; Sequence 4, Application US/095900020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 4
; TYPE: PRF
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match 9.7%; Score 281; DB 4; Length 623;
Best Local Similarity 26.9%; Pred. No. 1.8e-20;
Matches 90; Conservative 52; Mismatches 132; Indels 60; Gaps 12;

QY 67 SPRNSPGEVLLNLELGPENIPYLAHARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIIPP 126
DB 134 STSETQGGKILFKQSLGPLNLYSLAKLYVAVSGSVVEFRSISGSGVFGKLAIVVPP 193
QY 127 NF-PIDNLSAAQITWCPHVIVDVRQLEPNLPMEDVRNPFHYNCGSRLRLIAMLYTP 185
DB 194 GIEPVQSTSMLOQ---PHVLFARQVEPVIFAIPDLRSNLYHLMSDITDT-TSLVIMVYND 249
QY 186 LRANNSGDDVETVSC--RVLTRPSPDFSNFLVPTVBESKTKPFTLPILTISEMSNSRFP 243
DB 250 L-INPYANTNMSGGIVVEYKPGDFKHLKPPG-----SMLTHGSVPDILP 298
QY 244 ----VPIESLHTSPTEINIVVQNGRVTLDGELMGTTLQSPQICAFRGVLTSTRSRASD 299
DB 299 KSSSLWIGNRHSWDTDFIIR-----PFVQANRHFDFNQE 334
QY 300 QADTATPRLFNYYWHVQDLNLTGTPVDAED--IPG-PLGTDPDR-----GKVFVGAS 349
DB 335 TAGWSTPRFRPITITVSESNMKSGLIGVATDVIIVGIPDGNPDTTIPEQLTPAGIYSITA 394
QY 350 QRNLDSGTR----AHEAKVDTTAGRTPKLGSLEISTDSDDDQNOPTKFTPVGIGVDNE 405
DB 395 SNGTDTITTAAGYDAAETIVNTTNFKSMYICGSLQRAWDDKKISN---TAFITTAVRKGN 451

us-09-926-799-6-rai

57 VOAPGGETVSPRNSPGEVLLNLELGPENIPYLAHARMYNGYAGGFEVQVVLGNAFTA 116
61 VOAPGGETVSPRNSPGEVLLNLELGPENIPYLAHARMYNGYAGGFEVQVVLGNAFTA 120
117 GKIIIFAAIPPNFPIIDNLAAQITWCPHVIVDVRQLEPNLPMEDVRNPFHYNCGSRL 176
121 GKIIIVSCIPPFGSHNLITQAATLPHVIADVRLDPIEVLDPEDVRNVLFFHNDRNQTM 180
177 RLIAMLYDPLRANNSGDDVETVSCRVLTSPDPSDFSNFLVPTVBESKTKPFTLPILTISE 236
181 RLVCMLYDPLRANNSGDDVETVSCRVLTSPDPSDFSNFLVPTVBESKTKPFTLPILTISE 240
237 MNSRPPVPIESLHTSPTEINIVVQNGRVTLDGELMGTTLQSPQICAFRGVLTSTRSR 296
241 LNSRAPLPISSMGISPDNVQSVQFQNGRCTLDGLRGVLTTPVSLSHVAKIRG-----TSN 295
297 ASDQADTATPRLFNYYWHVQDLNLTGTPVDAEDIPGLGTDPDR-----KYFVGAS 349
296 GT-----VINLTDLGTPHPPFEG-PAPIGFDPDGGCDWHIINMTQFHSS 339
350 QRNLDSGTRAGRTPKLGSLEISTDSDDDQNOPTKFTPVGIGVDN----- 404
340 QTOYD-----VDITPDTFPHLGSIQAN-----GIGSGNVVGYL 373
405 -----EABFQOWSLPDYSGQFTNNMNLAPAVAPNPPGEGQLLFFRSQLPSSGGRSN 454
374 SWISPPSHPSGSQVLDWKINYGSSITEATHLAPSVVPPGFEVLVFFMSPKMPGQAYN- 432
455 GVLDCVLPQWQVHFOESAPQTOVALVRYNPDTKVLFEAKLHKLGMFTIANGDS- 513
433 --LPCLLPQBYISHLASEQAPTVEAALLHYDPDGTGRNLGEFKAYPDGFLICVPNGASS 490
514 -PITVPNGYFRPESWNPVTLAPMGTGNRR 545
491 GPQOLPINGVVFVSWRSRFLKPVGTASSAR 523

RESULT 2
US-09-590-020-6
; Sequence 6, Application US/095900020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilnis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRF
; ORGANISM: Feline calicivirus
US-09-590-020-6

Query Match 9.9%; Score 285.5; DB 4; Length 622;
Best Local Similarity 26.5%; Pred. No. 6e-21;
Matches 96; Conservative 52; Mismatches 151; Indels 63; Gaps 13;

QY 67 SPRNSPGEVLLNLELGPENIPYLAHARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIIPP 126
DB 134 STSETQGGKILFKQSLGPLNLYSLAKLYVAVSGSVVEFRSISGSGVFGKLAIVVPP 193
QY 127 NF-PIDNLSAAQITWCPHVIVDVRQLEPNLPMEDVRNPFHYNCGSRLRLIAMLYTP 185
DB 194 GIEPVQSTSMLOQ---PHVLFARQVEPVIFAIPDLRSNLYHLMSDITDT-TSLVIMVYND 249
```

RESULT 4  
US-09-590-020-7  
; Sequence 7, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-7

Query Match 9.7%; Score 279.5; DB 4; Length 626;  
Best Local Similarity 22.1%; Pred. No. 2.6e-20;  
Matches 122; Conservative 70; Mismatches 175; Indels 185; Gaps 22;

QY 67 SPRNSPGEVLLNLELGPENPVLHARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIAPP 126  
DB 134 STSETQGGKILFKQSLGLOPLNPLSHLAKLYVWSSGIEVRSISGSGVGGKLAIVVPP 193  
QY 127 NFPIDNLSAAQITMCPHVIVDVRQLEPNLPMVDRNNPFFHYNQSDSRLRLIAMLYTPL 186  
DB 194 G--VDPVQSTMLQYPHVLFDAQVEPVIFCLPDLRSLTYLHMSDSTD--TSLVIMYNDL 250  
QY 187 RANNGDDVFTVSC--RVLTRSPDFSNFLVPP-----TVESK----- 223  
DB 251 -INPYANDTSSGGCIIVTVETKPGDPFKPGLKPPGSMTHGSPVSDLPKSSSLWIGNRH 309  
QY 224 -----TKPPTL-----PI-LTISEMSNSRFPVPIESLHTSP 253  
DB 310 WSDITDFIIRPVFQANRHFENQTAGWSTFRPRPISVITTEQNGAKLGIV-----A 363  
QY 254 TENIVVQCGNG--RVTLDELGMGTTLQLPSQICAFRGVLTSTSTSRASDAQDTATPRLFNY 311  
DB 364 TDYIVPGIPDGWPDITIPGEL-----IPAGDYAITNGTN----- 398  
QY 312 YHHVQDLNLTNGTPYDPAEDIPGLGTPDPRGKVGVSQORNLDSITRAHEAKVDTTAGRP 371  
DB 399 -----DITATGYDTADIK--NNTNFRGMYICGSLQRAWG-----DKKISNTAFIT 443  
QY 372 TPKLGSLEISTDSD-----FDQOPTKFTPGVIGVDNEAEFQWMSLPDYSQGFTH 422  
DB 444 TATL-----DDNNKINPCNTIDQSKIIVVQDAHVGKKAQTSDDTLALLGYTG----- 492  
QY 423 NNNLAPAVANPFGQQL-----FFRSQLPSSGGRSNGVLDCLVPQEWVQH--FYQBS 473  
DB 493 -----IGEAIGSDRDRVVRISTLTPETGARGG-----NHPIFYKNS 528  
QY 474 AP-----AQTVALLRVYNP-----DTGKVLFEAKLHLKFWMT 506  
DB 529 IKLGIVIRSIDVNSQILHTSRQSLNHYLLPPDSFAVYRIIDNSGWSFIDIGDSGFSF 588  
QY 507 IANNGDSPITVP 518  
DB 589 VGVSGFGKLEFP 600

RESULT 5  
US-09-617-594A-4  
; Sequence 4, Application US/09617594A  
; Patent No. 6541458  
; GENERAL INFORMATION:

; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V  
; FILE REFERENCE: 454313-3151.1  
; CURRENT APPLICATION NUMBER: US/09/617,594A  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 668  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-617-594A-4

Query Match 9.7%; Score 279.5; DB 4; Length 668;  
Best Local Similarity 24.7%; Pred. No. 2.9e-20;  
Matches 94; Conservative 45; Mismatches 108; Indels 133; Gaps 16;

QY 67 SPRNSPGEVLLNLELGPENPVLHARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIAPP 126  
DB 179 STSETQGGKILFKQSLGLOPLNPLSHLAKLYVWSSGIEVRSISGSGVGGKLAIVVPP 238  
QY 127 NFPIDNLSAAQITMCPHVIVDVRQLEPNLPMVDRNNPFFHYNQSDSRLRLIAMLYTPL 186  
DB 239 G--IDPVQSTMLQYPHVLFDAQVEPVIFIPDLRSLTYLHMSDSTD--TSLVIMYNDL 295  
QY 187 -----RANNGDDVFTVSCRVLTSPDFSNFLVPPVPPVESKTKPFTLPTISEMSNSR 241  
DB 296 INPYANDTSSGGCIIVTVETKPGDPFKPGLKPP----- 328  
QY 242 FVPVIESLHTSFTENIVVQCGRVTLDELGMGTTLQLPSQICAFRGVLTSTSTSRASDOA 301  
DB 329 -----GSMLTGHSI--PSDLIPKSSSLWIG----- 351  
QY 302 DTATPRLFNYHHVQDLNLTNGTPYDPAEDIPGLGTPDPRGKVGVSQORNLDSITRAHE 361  
DB 352 -----NRHW-----SDI-----TDFVIKPFVFOANRHFDFN----- 377  
QY 362 AKVDTTAG-----RFTP-----KLSLEISTDS-----DDP-DQOPTKFTVPV 399  
DB 378 --QETAGWSTFRPRPITTVSEKSGSKLG-IGVATDSIVPGIPDGWPDITTIPEKLTPEG 433  
QY 400 -IGVDNEAEFQWMSLPDYS 418  
DB 434 DYAITNGGNNDITTAADYDG 453

RESULT 6  
US-09-590-020-2  
; Sequence 2, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Feline calicivirus

US-09-590-020-2

Query Match 9.5%; Score 276; DB 4; Length 623;  
Best Local Similarity 26.2%; Pred. No. 6e-20; Mismatches 164; Indels 54; Gaps 16;  
Matches 100; Conservative 63;

QY 67 SPRNPGVLLNLELPEINPYLAHARMYNGYAGGFVQVVLNAGNATGKIIFAIAPP 126  
DB 134 STSETQKILFKQSLGPLNLYSLAKLYVAMSGSVREFSISGSGVFGKLAIVVPP 193

QY 127 NF-PIDNLSAAQITMCPHVIVDVRQLEPNLMPDVRNFFHYNGSDSRLRLIAMLYTP 185  
DB 194 GIEPVQSTSMLOQ---PHVLFARQVPEVIFAIPDLRSLNLYHLMSTDT-TSLVIMVYND 249

QY 186 LRANNSGDDVFTVSC--RVLTRPSDFSENFVPPTVESKTKPFTLPILITISEMSNSRFP 243  
DB 250 L-INPVANDTNSGGCVIVETKPGDPFKHLLKPGF-----SMLTHGSVPSDLIP 298

QY 244 ---VPISLHTSPTENIVVQ-----CONGRVTLDGELMG--TTQLLPQICAFRGVLTFR- 292  
DB 299 KSSSLWIGNRHWSDDITDFIIRPVFQANRHFDFNQETAGWSTPRFRPITIVSESNMSKL 358

QY 293 STSRASDQ-----ADTATPRLEFNYWHVOLDNLGT-----PYDPAEDIPGLGT 337  
DB 359 GIGVATDYIVPGIDGWPDPITPEQUTPAGIYSITASNGIVITTAAGYDAETI---VNT 415

QY 338 PDFRGKVFVGASQRNLDSSTTRAHEAKVDTTAGRTTPKLGSLSEISTDSDDFQDQNTKFTTP 397  
DB 416 TNFKSMYICSLQRANGDKKISNTAFI-TTAVR---KGNISIE---PSNTIDMTKLIVYQD 468

QY 398 VGIGVDNEAFQOWSLPDYSG 418  
DB 469 AHVGEVQTSITLALLGYTG 489

RESULT 7  
US-09-617-594A-2  
; Sequence 2, Application US/09617594A  
; Patent No. 6541458  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.1  
; CURRENT APPLICATION NUMBER: US/09/617,594A  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 669  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-617-594A-2

Query Match 9.0%; Score 262; DB 4; Length 669;  
Best Local Similarity 22.3%; Pred. No. 2.1e-18;  
Matches 110; Conservative 69; Mismatches 166; Indels 148; Gaps 20;

QY 30 DPVAGAAIAAPTITGQNIIDPHIMNFVQAPGGETV-----SPRNSPGEVLLNLELGP 83  
DB 146 EPSAQMATAADAATGKSDSEW--ESF-----SFHTSVNWSVSETQKILFKQSLGP 196

QY 84 ENPYLAHARMYNGYAGGFVQVVLNAGNATGKIIFAIAPPNPFIDNLSAAQITMCPH 143  
DB 197 LNPYLEHLSKLYVAMSGVDVRFISGSGVFGKLAIVVPPG--VDPVQSTSMLOQPH 254

QY 144 VIVDVRQLEFPVNLMPDVRNFFHYNGSDSRLRLIAMLYTPI-----RANNSGDDVFTV 198

DB 255 VLFDAQRQVEPVIFISIPDLRSTLYHLMSTDT-TSLVIMVYNDLINPYANDSNSGCIIVTV 313

QY 199 SCRVLTRPSDFSENFVLPVP-----TWESKTKPFTL 229

DB 314 E-----TKPGDPFKHLLKPGSMMLTHGSIPLSDLPKSSSLWIGNRYWSDITDFVIRPEVF 369

QY 230 -----PI-LTISEMSNSRFPVIESLHTSPTENIVVQCONG-RV 266

DB 370 QANRHFDFNQETAGWSTPRFRPITITISESNGSLGTGV-----ATDYIVFPIGIDGWP 423

QY 267 TLDGELMGTTQLLPQICAFRGVLTFRSISRASDQADTATPRLFNYYWHVOLDNLNGTPVD 326

DB 424 TTIGE-----ELTFA-----GDYSITNGSGNDIA-----TANAYD 453

QY 327 PAEDIPGLGTDPFRGKVFVGASQRNLDSSTTRAHEAKVDTTAGRTTPKLGSLSEISTDSDD 386

DB 454 SADVI---TNTTFRGMWYICGALQRAWG-----DKKISSTA-FITTAIKEGNTLKPSNT 503

QY 387 FDNQNTKFTTPVGIGVDNEAFQOWSLPDYSGOFTTHNMNLAPAVAPNPFGEQLLFFR--- 443

DB 504 IDMTKIAVYQDTHVGRDVQTSDDTLAILGYTG-----IGBQAIGSNRDS 547

QY 444 ----SOLPSSGGR 452

DB 548 VVRISMLPETGAR 560

RESULT 8  
US-09-134-001C-4463  
; Sequence 4463, Application US/09134001C  
; Patent No. 6380370  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC-007  
; CURRENT APPLICATION NUMBER: US/09/134,001C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/064,964  
; PRIOR FILING DATE: 1997-11-08  
; PRIOR APPLICATION NUMBER: US 60/055,779  
; PRIOR FILING DATE: 1997-08-14  
; NUMBER OF SEQ ID NOS: 5674  
; SEQ ID NO 4463  
; LENGTH: 2137  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-09-134-001C-4463

Query Match 4.4%; Score 128.5; DB 4; Length 2137;  
Best Local Similarity 23.8%; Pred. No. 0.0021;  
Matches 86; Conservative 40; Mismatches 139; Indels 97; Gaps 18;

QY 219 TWESKTKPFTLPILTISEMSNSRFPVIESLHT-----SPT-----E 255

DB 243 TSEARTSTNQISNLITSTSNQSSPTSFANLRTFSRFTVLTNTMAAPTITTTTSSLSRN 302

QY 256 NIVVQCON-----GRVTLDELMGTQLLPQICAFRGVLTFRSISRASDQ--DTAT 305

DB 303 SVVVKNDNFNEHMLNLSGSATYDPK-TGIATLTPEAY-SQKGAISLNLRLDSNRSFRFTCK 360

QY 306 PRLFNYYWHVOLDNLNGTPYDPAEDIPGLGTPDFRGKVFVGASQRNLDSSTTRAHEAKVD 365

DB 361 VNUGNRYEGSPDVTGGDGGIGFAPSPGLQIGKEGAAGVIGGLNN-----AFGFKLD 414

QY 366 TTAGRETPPK-----LGSLEISTDSDDFQDQNTKFTPVGI-GVDNEAE 407

DB 415 TYHTSTPKSDAKAKADPRNVGGGAFGAF-VSTDRNGMATTEASAAKLNQPTDN--S 471

QY 408 FQOWSLPDYSGQ-----FTHNMNLAPAVAPNPFGEQLLFFRSQLPSSGSRNG 455

DB 472 FQDFVI-DYNGDITKVMVTYVAGQTFTRNLT---DWIKNSGG--TTFSLSWTASTGAKN- 524

QY 456 VLDCLVPOEWHQYQESAPAOQVALVRYVNPDTGKVLFEAKHLKLGPMWTIANNGDSPI 515  
Db 525 -----LQOVQGTPTXTESA-----VAKVRYVDANTGKIIPPK-----TIAGEVDATV 568  
QY 516 TV 517  
Db 569 NI 570

RESULT 9  
US-08-961-083-86  
; Sequence 86, Application US/08961083  
; Patent No. 6159469  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/961,083  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER:  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 86:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 776 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-961-083-86

Query Match 4.2%; Score 121; DB 3; Length 776;  
Best Local Similarity 20.5%; Pred. No. 0.0023;  
Matches 110; Conservative 75; Mismatches 233; Indels 118; Gaps 25;

QY 31 PVAGAAIAAPLTGQCOQNIIDPWIMNFVQAPGGEFTVSPRNSPGEVLLNLELGPINPYLA 90  
Db 58 PLAFASDSNPDPVSNVDKLI SYN--NQPANRWTNNRTP--EASVGLFGD----- 107  
QY 91 HLARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIPNPFIDNLSAAQITMCP----- 142  
Db 108 -----SGILSKRSVDNLSVGFHEDHG-----VGVPKSYVIEYVGVKTPTAPKNPSFVGN 157  
QY 143 --HVIYDVRLQLEPV-NLPWPDV-----RNPFHYNQSDSLRL-----IA 180  
Db 158 EDHVFNDSANWKEPVTLKAPQLKAGEMNHFSDFKVETAYVIRVMYKADNKRGT SITEVQ 217  
QY 181 MLYTPLRANNSGDVFTVSCRVLTRPSDPFSNFLVPPTVESKTKPFTLPILITIGEMNS 240  
Db 218 IFAKQVAAKQGTQRLQVQDKLANPDLTDYLL--ESVDGK-----VPATAS-VSN 269  
QY 241 RFPVPIESLHTSPTENIVVQCQNGRVTLDGELMGTTQLLPQLCAFRGLVTRSTRASDQ 300

Db 270 GLATVVPVSREGEPRVRIAKAEN-----GDILGBYRL-----HFTKDKSLLSHKPVAANKQ 320  
QY 301 A-----DTATPRLENYIMHVQLDNLNGTYPD--PAEDIPGPGPLGTDPDFRGKVFVGA 348  
Db 321 ARLLQVQGALELPTKVPVYFTGKCGYETKOLTVHEEVPAENLT-KAGQFTVGRVLG-- 377  
QY 349 SORNLDSTTRAHEAKVDTTAGFTPKLASLEISTSDDFDQNOPTKFTPVGIGVDNEAEF 408  
Db 378 --SNL-----VAEITVRVTDKLG--ETLSDPNPYDENSNAQFASATNDIDKNSHD 423  
QY 409 QQWSLPDYSGQFTHN---MNLAPAVAPNPPGQQLPERSQLPSSSGRSNG-VLDCILVPOE 464  
Db 424 RVDYLDN--GDHSENRRTNWSPTSPSNPEVSAGVIP-----RENGKIVERTVTQ 472  
QY 465 WYQHIFYQESAPAQTOQVALVRYVNPDTGKVLFEAKHLKLGFTIANNGDSPI TVPPN 520  
Db 473 KVQFPADSGTDAPSKVLERYVGPB-----FVPTYSNYQ--AYDAHPPNPEN 521

RESULT 10  
US-09-536-784-86  
; Sequence 86, Application US/09536784  
; Patent No. 6573082  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/536,784  
; FILING DATE: 30-Oct-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: OCT-30-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Michelle S. Marks  
; REGISTRATION NUMBER: 41,971  
; REFERENCE/DOCKET NUMBER: PB340P3  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 86:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 776 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 86:  
US-09-536-784-86

Query Match 4.2%; Score 121; DB 4; Length 776;  
Best Local Similarity 20.5%; Pred. No. 0.0023;  
Matches 110; Conservative 75; Mismatches 233; Indels 118; Gaps 25;

QY 31 PVAGAAIAAPLTGQCOQNIIDPWIMNFVQAPGGEFTVSPRNSPGEVLLNLELGPINPYLA 90  
Db 58 PLAFASDSNPDPVSNVDKLI SYN--NQPANRWTNNRTP--EASVGLFGD----- 107  
QY 91 HLARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIPNPFIDNLSAAQITMCP----- 142

Db	108	-----SGILSKRSVDNLISVGFHEDHG-----VGVPEKSYVIEYYVGKVTPTAPKNPSFVGN	157
Qy	143	-----HVIVDVROLEPV-NLMPDV-----RNNFFHYNOGSDSRRL-----IA	180
Db	158	EDHVFNDSANWKPVTNLKAPQAKAGEMNHFSFDKVEYAVIRVMKADNKRGTSITEVQ	217
Qy	181	MLYTPLRANNSGDVFTVSCRVLTRPSPDFSNFLVPPVTESKTKPFTLPILITISEMSNS	240
Db	218	IFAKQVAAAKQGTQIQVDGKDLANFNPLDLYL--ESVDGK-----VPAVTAS-VSNN	269
Qy	241	RPVPIESLHTSTENIVVQCONGRVTLDELGMGTQLLPSQICAFRGVLTRSTSRASDQ	300
Db	270	GLATVVPVSREGEPRVIAKEN-----GDILGEYRL-----HFTKDKLSLHKPVAAYKQ	320
Qy	301	A-----DTATPRLFNYWHVQLDNLNGTPYD-PAEDIPGLGTGPPDFRGKVFVA	348
Db	321	ARLLQVQGALELPKVPVFTGKGVEYTKDLIVEMEVEVFAENLT-KAQQFTVRGRVLG--	377
Qy	349	SORNLSTTRAHBAKVDYTAGRTPKGLSLEISTDSDPDQNPQKFTFPVGVGVNDNEAF	408
Db	378	--SNL-----VAEITVRVTDKLG--ETLSDNPNYDENSNOAFASATNDIDKNSHD	423
Qy	409	QOWSLPDYSQFTHN--WNLAFAVAPNPFGEQLLFFRSQPSGGRSNG-VLDCILVPOE	464
Db	424	RDYLAND--GDHSENRRTNWSPTPSNPEVSAGVIP-----RENGKIVERTVIGQ	472
Qy	465	WVQHFOQESAPAOQVALVRYVNPDTGKVLFEAKLHKLGFMTIANNQDSPIVPPN	520
Db	473	KVQFFADSGTDAKSLVLERYVGPE-----FEVPTVYSNYQ--AYDADHPFNPNPEN	521
RESULT 11			
US-09-091-219-24			
; Sequence 24, Application US/09091219			
; Patent No. 6171592			
; GENERAL INFORMATION:			
; APPLICANT: STUDDERT, Michael J.			
; APPLICANT: CRABB, Brendan S.			
; APPLICANT: FENG, Li			
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS			
; FILE REFERENCE: 040268/0151			
; CURRENT APPLICATION NUMBER: US/09/091,219			
; CURRENT FILING DATE: 1998-10-05			
; EARLIER APPLICATION NUMBER: PCT/AU96/00815			
; EARLIER FILING DATE: 1996-12-18			
; EARLIER APPLICATION NUMBER: AU PN7201			
; EARLIER FILING DATE: 1995-12-18			
; NUMBER OF SEQ ID NOS: 25			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 24			
; LENGTH: 2318			
; TYPE: PRT			
; ORGANISM: Foot-and-mouth disease virus			
US-09-091-219-24			
Query Match			
Best Local Similarity 4.1%; Score 119; DB 3; Length 2318;			
Matches 106; Conservative 60; Mismatches 221; Indels 156; Gaps 24;			
Qy	79	LELGEINPYLAHLARMYNGVAGFEVQVVLGNAFTAGKIIFAAIPNFPIDNLSAAQI	138
Db	367	LELPTDHKGVYGLTDSYAYMNGWDVEVTAVGNQFNGCGLLVAMVPELYSIQRELYQL	426
Qy	139	TWCPHVIVDVRQLEPNLMPDVRNFFHYNGQSDSR-LRLIAMLTYPLRANNSG----	192
Db	427	TLFPHQFINPRNTMHAITVPFVGN--RYDQYKVHKPWLTVMVVAELTVNTEGAPOIK	484
Qy	193	-----DDVFTVSCR-----VLRP-----SPDFSFNFLVP-----	217
Db	485	VYANIAPTNNHVAGPEPSKEGIFPVACSDGYGGLVTTDPKADPVYGVKFNPPRNLPCR	544
Qy	218	-----PTVESKT-----KPFTLPILITISEMSNSRFPVIESLH	250
Db	545	FTNLLDVAEACPTFLRPEGVYVTKTDSRVLAQFDMNS-LAAQKMSNT-FLAGLAQY	602
Qy	251	TSPTENIVQCONGRVTLDELGMGTQLLPSQICAFRGVLTRSTSRASDAQATATPRLFN	310
RESULT 12			
US-09-660-541-24			
; Sequence 24, Application US/09660541			
; Patent No. 6531136			
; GENERAL INFORMATION:			
; APPLICANT: STUDDERT, Michael J.			
; APPLICANT: CRABB, Brendan S.			
; APPLICANT: FENG, Li			
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS			
; FILE REFERENCE: 040268/0151			
; CURRENT APPLICATION NUMBER: US/09/660,541			
; CURRENT FILING DATE: 2000-09-12			
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/091,219			
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05			
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201			
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18			
; NUMBER OF SEQ ID NOS: 25			
; SOFTWARE: PatentIn Ver. 2.0			
; SEQ ID NO 24			
; LENGTH: 2318			
; TYPE: PRT			
; ORGANISM: Foot-and-mouth disease virus			
US-09-660-541-24			
Query Match			
Best Local Similarity 4.1%; Score 119; DB 4; Length 2318;			
Matches 106; Conservative 60; Mismatches 221; Indels 156; Gaps 24;			
Qy	79	LELGEINPYLAHLARMYNGVAGFEVQVVLGNAFTAGKIIFAAIPNFPIDNLSAAQI	138
Db	367	LELPTDHKGVYGLTDSYAYMNGWDVEVTAVGNQFNGCGLLVAMVPELYSIQRELYQL	426
Qy	139	TWCPHVIVDVRQLEPNLMPDVRNFFHYNGQSDSR-LRLIAMLTYPLRANNSG----	192
Db	427	TLFPHQFINPRNTMHAITVPFVGN--RYDQYKVHKPWLTVMVVAELTVNTEGAPOIK	484
Qy	193	-----DDVFTVSCR-----VLRP-----SPDFSFNFLVP-----	217
Db	485	VYANIAPTNNHVAGPEPSKEGIFPVACSDGYGGLVTTDPKADPVYGVKFNPPRNLPCR	544
Qy	218	-----PTVESKT-----KPFTLPILITISEMSNSRFPVIESLH	250
Db	545	FTNLLDVAEACPTFLRPEGVYVTKTDSRVLAQFDMNS-LAAQKMSNT-FLAGLAQY	602
Qy	251	TSPTENIVQCONGRVTLDELGMGTQLLPSQICAFRGVLTRSTSRASDAQATATPRLFN	310



```

; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEFAX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-503-2

Query Match 4.0%; Score 117; DB 2; Length 2206;
Best Local Similarity 18.6%; Pred. No. 0.036;
Matches 121; Conservative 71; Mismatches 197; Indels 260; Gaps 28;

QY 11 SNDGAAGLVPEINNEAMALDPA-----GAATAAPLT----- 42
Db 270 TNSATIVLPYVN--ALAIIDSMVKHNNWGIALPLSLDPAQDSSVEIPITVTIAPMCSE 327
QY 43 --GQNIIDP-----WIMNFVQAPGG--EFTVSPR-NSPGEV----- 75
Db 328 ENGLRNVTAPEQGLPVLTGSGNQYLTSNQHSPCAIPEFDVTPPIDIPGEVKNMELA 387
QY 76 -----LNLGLGPEINPYLAH-----LARMY 96
Db 388 EIDTMIPLNLESTKRNMTDMYRVLTSDSADLSQPICLSLSPAFDRLSHTMLGEVLNY 447
QY 97 NGYAGGFEVQVVLGNAGFTAGKIIFAAIPPNFPIDNLSAAQITMCPHVIVDVRLQLEPVNL 156
Db 448 THWAGSLKFTFLFCGSMMATGKILVAYAPPGAQ--PPTSRRKEAMLGTHVINDLGLQSSCTM 506
QY 157 PMPDVRNFFHYNQS--DS-----RLRLIAMLYTPLRANSGDDVFTVSCRVL 204
QY 507 VVPWISN--VTYRQTQDSFTEGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN--- 558
Db 507 VVPWISN--VTYRQTQDSFTEGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN--- 558
QY 205 RPSDFSFNFI-----VPTVESKTKPFTLPILTIS-----EIMSRSFPVPPIES 248
Db 559 ----DFSVLLRDTTHISQSALPQGIEDLTSEVAQGALTLSPKQDQSLPDTTKASGPAHS 614
QY 249 LHTSPTENIVVQCNGRVTLTGELMGTTQLLPQICAFRGVL--TRSTSRASDAQDATPR 307
Db 615 KEV-PALTAV-----ETGATNPLAPSDTVQTRHVVRQRSESTIESFFA--- 658
QY 308 LFNYWHVQDLNLTGTPYDPAEDIPGLTDPFGKVGFGVASQRNLDSTTRAHEAKVDIT 367
Db 659 -----RGACVAIIIEVDNEQPTTRAQ----- 678
QY 368 AGRFTPKLSLEISTDSDPDQONQTKFTPVGIGVDNEAEFQWLSLPDYSGQFTH---NM 424
Db 679 -----KLFAMWRITKYDVTQLRKLEF-----FTYSRFD 708
QY 425 NLAPAVAPNPPGQLLFFRSQPSGGRSNGVLDCL-----VPQEWVQHFYQES---- 473
Db 709 EFTFVVTANFTN-----ANNGHALNQVYQIMYIPPGAPTPKSWDDYTWQTSNPS 758
QY 474 -----APAQVQVALRVNPDQKVLFEAKLHLGFWTIANN--GDS 513
Db 759 IFYTYGAAPARISVPYVGLANAYSH---FYDGFAGVAPLKTANDQIGDS 804

RESULT 15
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racanietello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
;

CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release 1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,250
FILING DATE: 6-JUN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 278-0400
TELEFAX: (212) 391-0525
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2206 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-465-250-2

Query Match 4.0%; Score 117; DB 3; Length 2206;
Best Local Similarity 18.6%; Pred. No. 0.036;
Matches 121; Conservative 71; Mismatches 197; Indels 260; Gaps 28;

QY 11 SNDGAAGLVPEINNEAMALDPA-----GAATAAPLT----- 42
Db 270 TNSATIVLPYVN--ALAIIDSMVKHNNWGIALPLSLDPAQDSSVEIPITVTIAPMCSE 327
QY 43 --GQNIIDP-----WIMNFVQAPGG--EFTVSPR-NSPGEV----- 75
Db 328 ENGLRNVTAPEQGLPVLTGSGNQYLTSNQHSPCAIPEFDVTPPIDIPGEVKNMELA 387
QY 76 -----LNLGLGPEINPYLAH-----LARMY 96
Db 388 EIDTMIPLNLESTKRNMTDMYRVLTSDSADLSQPICLSLSPAFDRLSHTMLGEVLNY 447
QY 97 NGYAGGFEVQVVLGNAGFTAGKIIFAAIPPNFPIDNLSAAQITMCPHVIVDVRLQLEPVNL 156
Db 448 THWAGSLKFTFLFCGSMMATGKILVAYAPPGAQ--PPTSRRKEAMLGTHVINDLGLQSSCTM 506
QY 157 PMPDVRNFFHYNQS--DS-----RLRLIAMLYTPLRANSGDDVFTVSCRVL 204
QY 507 VVPWISN--VTYRQTQDSFTEGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN--- 558
Db 507 VVPWISN--VTYRQTQDSFTEGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN--- 558
QY 205 RPSDFSFNFI-----VPTVESKTKPFTLPILTIS-----EIMSRSFPVPPIES 248
Db 559 ----DFSVLLRDTTHISQSALPQGIEDLTSEVAQGALTLSPKQDQSLPDTTKASGPAHS 614
QY 249 LHTSPTENIVVQCNGRVTLTGELMGTTQLLPQICAFRGVL--TRSTSRASDAQDATPR 307
Db 615 KEV-PALTAV-----ETGATNPLAPSDTVQTRHVVRQRSESTIESFFA--- 658
QY 308 LFNYWHVQDLNLTGTPYDPAEDIPGLTDPFGKVGFGVASQRNLDSTTRAHEAKVDIT 367
Db 659 -----RGACVAIIIEVDNEQPTTRAQ----- 678
QY 368 AGRFTPKLSLEISTDSDPDQONQTKFTPVGIGVDNEAEFQWLSLPDYSGQFTH---NM 424
Db 679 -----KLFAMWRITKYDVTQLRKLEF-----FTYSRFD 708
QY 425 NLAPAVAPNPPGQLLFFRSQPSGGRSNGVLDCL-----VPQEWVQHFYQES---- 473
Db 709 EFTFVVTANFTN-----ANNGHALNQVYQIMYIPPGAPTPKSWDDYTWQTSNPS 758
QY 474 -----APAQVQVALRVNPDQKVLFEAKLHLGFWTIANN--GDS 513
Db 759 IFYTYGAAPARISVPYVGLANAYSH---FYDGFAGVAPLKTANDQIGDS 804
```





**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 35.0315 Seconds  
(without alignments)  
4368.312 Million cell updates/sec

Title: US-09-926-799-6  
Perfect score: 2896  
Sequence: 1 MKMASNDAPSDGAGLVP.....VNPFYTLAPMTGTGRRRIQ 548

Scoring table: BLOSUM62

Gapex 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

1:	/cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
2:	/cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
3:	/cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
4:	/cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
5:	/cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
6:	/cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
7:	/cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
8:	/cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
9:	/cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
10:	/cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
11:	/cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
12:	/cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
13:	/cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
14:	/cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
15:	/cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
16:	/cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
17:	/cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
18:	/cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1211	41.8	530	14	US-10-314-739-3
2	279.5	9.7	668	14	US-10-209-507-4
3	275.5	9.5	547	12	US-10-670-695-4
4	275.5	9.5	671	12	US-10-670-695-2
5	262	9.0	669	14	US-10-209-507-2
6	126	4.4	1965	15	US-10-369-493-2646
7	121.5	4.2	522	14	US-10-172-502-17
8	121.5	4.2	1831	12	US-10-282-122A-71033
9	121	4.2	338	10	US-09-867-932A-2
10	121	4.2	776	9	US-09-765-272-86
11	121	4.2	2233	10	US-09-769-787-2
12	116.5	4.0	1147	15	US-10-327-481A-38
13	112.5	3.9	932	14	US-10-174-677-88
14	112	3.9	709	12	US-10-231-956A-287
15	111.5	3.9	485	14	US-10-172-502-18

16	111.5	3.9	2271	12	US-10-282-122A-43924
17	111.5	3.9	2283	14	US-10-172-502-4
18	111	3.8	5175	15	US-10-120-801-74
19	111	3.8	5175	15	US-10-369-493-6859
20	111	3.8	5175	15	US-10-369-493-6861
21	111	3.8	5198	12	US-10-346-863-34
22	111	3.8	5198	12	US-10-120-801-75
23	111	3.8	5198	15	US-10-369-493-6858
24	111	3.8	5198	15	US-10-369-493-6860
25	110.5	3.8	5877	14	US-10-142-515-11
26	110.5	3.8	5935	14	US-10-243-243A-8
27	109.5	3.8	2344	9	US-09-815-242-12713
28	109	3.8	6310	12	US-10-282-122A-67793
29	109	3.8	19662	15	US-10-084-846A-6
30	108.5	3.7	997	15	US-10-369-493-3593
31	108.5	3.7	2005	9	US-09-735-367B-3
32	108.5	3.7	2063	9	US-09-735-367B-2
33	107.5	3.7	487	15	US-10-369-493-22178
34	107	3.7	1686	15	US-10-386-414-2
35	107	3.7	2478	9	US-09-815-242-5816
36	107	3.7	2478	9	US-09-815-242-12967
37	106.5	3.7	1794	10	US-09-965-738-299
38	106.5	3.7	1799	10	US-09-965-738-149
39	106.5	3.7	1821	10	US-09-965-738-82
40	106.5	3.7	2234	12	US-10-622-090-20
41	106.5	3.7	2358	12	US-10-282-122A-45763
42	106.5	3.7	11721	10	US-09-965-738-162
43	106	3.7	920	15	US-10-320-797-3323
44	105.5	3.6	435	16	US-10-389-566-1432
45	105	3.6	1255	12	US-10-312-352-23

#### ALIGNMENTS

RESULT 1  
US-10-314-739-3  
; Sequence 3, Application US/10314739  
; Publication No. US20030129588A1  
; GENERAL INFORMATION:  
; APPLICANT: Estes, Mary K  
; Jiang, Xi  
; Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; Characterize No. US20030129588A1walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski L.L.P.  
; CITY: Washington, D.C.  
; STATE: <Unknown>  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/10/314,739  
; APPLICATION NUMBER: US/10/314,739  
; FILING DATE: 09-Dec-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX: <Unknown>

```

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match 41.8%; Score 1211; DB 14; Length 530;
Best Local Similarity 45.4%; Pred. No. 8.8e-108;
Matches 260; Conservative 74; Mismatches 161; Indels 78; Gaps 11;

QY 1 MKQASNDAPSDGAG---LYPEIN-NEAMALDPVAGAAIAPLTGQOIIDPWIMNF 56
Db 1 MMASDKATSSVDGASGAGQOLPEVNASDPLANDPVAGSTAVATAGQVNPDIWINF 60
QY 57 VQAPGCEFTVSPRNSGVEVLLNLELGPINPYLAHARMYNGYAGGFEVQVVLNAGETA 116
Db 61 VQAPQCEFTISPNPTGDLVLDLGLPHLNPFLHLSQMYNGWGNVRIMLAGNAFTA 120
QY 117 GKIIFAAIPNPFIDNLAAQITMCPHVIQVLEPNLMPDVRNNFFHYNQSDSL 176
Db 121 GKIIVCIPPGGSHNLTAQATLFPHVIADVTLPDIEVPLEDVNVLFNHNDRNQTM 180
QY 177 RLIALMYTPLRANNSGDDVFTVSCRVLTRPSDFSNFLVPPVTESKTPFTLPILTISE 236
Db 181 RLVCMLYTPLRGTGCTDSFVAGRVMTCPSPDFNLFVLPVPTVEQKTRPFTLPNPLSS 240
QY 237 MNSRRFPVPIESLHSTENIVQCGRVTLTGELMGITQLLPSQICAFRGVLTSTSR 296
Db 241 LNSRAFLPISSMGISSPDNVQSQVQFQNGRCTLGRLVGTTPVSLSHVAKIRG---TSN 295
QY 297 ASDQADATPLFLNYHVHQLDNLNGTPYDPAEDIPGLTPTDFRG-----KVFQVAS 349
Db 296 GT-----VINTELDGTFFHPFEG-PAPIGFDLGGCDWHNNMTQFCHSS 339
QY 350 QRNLDSITTRAHEAKVDTTAGFTPKLGLSLEISTDSDDFDQNTKFTPVGIVDN----- 404
Db 340 QTQYD-----VDTTPTDFVPHLGSIQAN-----GIGSGNVYGV 373
QY 405 -----EAFQWLSLPDYSQGFTHNMNLAPAVAPNFPGEQLLFFRSQLPSSGGRSN 454
Db 374 SWISPPHPSQDVLKIPNGSSITEATHLASVYPGPGFGEVLVFFMSKMPGPGAYN- 432
QY 455 GVLDCLVPOEWQHFYQESAPAQTOVALRVYVNDPTGKVLFEAKLHKLGFMTIANNGDS- 513
Db 433 --LPCLLPQEIYSHLASEQAFTVGEAALLHYVDPDTRNLGELFKAYPDGFLTCVPNGASS 490
QY 514 -PITVPNGYFRFESWNPFFYTLAPMGTGNGER 545
Db 491 GPQQLPBGVFFVSVSVSRFYQLKPVGTASSAR 523

RESULT 2
US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT FILING DATE: 2002-11-12
; PRIOR FILING DATE: 2002-11-12
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR FILING DATE: 2000-02-11
; PRIOR FILING DATE: 2000-02-11
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-209-507-4

Query Match 9.7%; Score 279.5; DB 14; Length 668;
Best Local Similarity 24.7%; Pred. No. 1.5e-17;
Matches 94; Conservative 45; Mismatches 108; Indels 133; Gaps 16;

QY 67 SPRNSGVEVLLNLELGPINPYLAHARMYNGYAGGFEVQVVLNAGNAFTAGKIIFAAIIPP 126
Db 179 STSETQKILFKQSLGPLNPLYLTHLAKLYAVAGSGIEVRFSGSGVFGKLAIVVPP 238
QY 127 NFPIDNLAAQITMCPHVIQVLEPNLMPDVRNNFFHYNQSDSLRLIAMLYTPL 186
Db 239 G--IDPVQSTSMQVPHVLFDAQVEPVFIPIPLDRLNSLYHLSMDTDT-TSLIVIMYNDL 295
QY 187 -----RANNSGDDVFTVSCRVLTRPSDFSNFLVPPVTESKTPFTLPILTISEMSNR 241
Db 296 INFYANDSNSGCIITVE---TKPGDPFKFLLKPP----- 328
QY 242 FVPVIESLHSTENIVQCGRVTLTGELMGITQLLPSQICAFRGVLTSTSRASDQA 301
Db 329 -----GSMUTHGSI--PSDLIPKSSSLWIG----- 351
QY 302 DTATPRLFNYYHVQLDNLNGTPYDPAEDIPGLTPTDPRGKVGVASORNLDSITTRAHE 361
Db 352 -----NRHW-----SDI-----TDFVIKPFVFOANRHFEN----- 377
QY 362 AKVDTTAG-----RPTP-----KLGSLISTDS-----DDF-DQNOPTKFTPVG 399
Db 378 ---QETAGWSTPRPTTITVSKGSGKLG-IGVATDSIVPGIPDGPMDPTTPEKLTPTAG 433
QY 400 -IGVDNEAEFQWLSLPDYS 418
Db 434 DYAITNGGNDITTAADYDG 453

RESULT 3
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR FILING DATE: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-670-695-4

Query Match 9.5%; Score 275.5; DB 12; Length 547;
Best Local Similarity 22.3%; Pred. No. 2.7e-17;
Matches 123; Conservative 70; Mismatches 174; Indels 185; Gaps 23;

QY 67 SPRNSGVEVLLNLELGPINPYLAHARMYNGYAGGFEVQVVLNAGNAFTAGKIIFAAIIPP 126
Db 55 STSETQKILFKQSLGPLNPLYLTHLAKLYAVAGSGIEVRFSGSGVFGKLAIVVPP 114
QY 127 NFPIDNLAAQITMCPHVIQVLEPNLMPDVRNNFFHYNQSDSLRLIAMLYTPL 186

```

Db 115 G--VDPVQSTSLMQYPHVLFDAQVEPVIFCPDLKSTLYHLMSDQDT--TSLVIMVYNDL 171  
Qy 187 RANNSGDDVFTVSC--RVLTSPSPDFSNFLVPP-----TVES-----KT-----224  
Db 172 -INPANDANSGCCIVVETKPGDPFKFHLKPPGSMTHGSIPLDLPKTSLSLWGNRY 230  
Qy 225 -----KPTTL-----PI-LTISEMNSRPPVPPIESLHTSP 253  
Db 231 WSDITDFVIRPVFOANRHDFNQETAGSTPRFRPISVTITEQNGAKLIGV-----A 284  
Qy 254 TENIVVQONG--RVTLDELGTTQLLPSQICAFRGVLTSTRSASDAQADTATPRLFNY 311  
Db 285 TDYIVPGIPDGWPDPTIPGEL-----IPAGDYAITNGTN-----319  
Qy 312 YHVHQLDNLNGTTPYPAEDI PGPLGTDPFRGVGVASORNLDSITTRAHEAKVDTTAGRF 371  
Db 320 -----DITATGYDTADIK--NNTNFRGMVYCGSLQRAWG-----DKISNTAFIT 364  
Qy 372 TPKGLSLEISTDSD--FDQOQTKFTPVGIGVDNEABFQWMSLPDYSQGFTH 422  
Db 365 TATL-----DGDNNKINPCNTIDQSKI VVQDNHVGKKAQTSDDTLALLGYTG----413  
Qy 423 NNNLAPAVAPNPFGEQLL-----FFRSQLPSSGSRNGVLDCLVPQEWVQH--FYQBS 473  
Db 414 -----IGEQAIGSDRDRVVRISTLPETGARG--NHPIFYKNS 449  
Qy 474 AP-----AQTOVALRVYVNP-----DTGKVLPEAKLHKLGMFT 506  
Db 450 IKLGVVIRSIDVFNSQLHTSRQLSLNHYLLPPDSFAVYRIIDNSGSMFIDIGSDGFSF 509  
Qy 507 IANNGDSPITVP 518  
Db 510 VGVSGFGKLEFP 521

## RESULT 4

US-10-670-695-2

; Sequence 2, Application US/10670695

; Publication No. US20040058316A1

; GENERAL INFORMATION:

; APPLICANT: Jensen, Wayne A.

; APPLICANT: Lappin, Michael R.

; APPLICANT: Rosen, David K.

; APPLICANT: Andrews, Janet S.

; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE

; TITLE OF INVENTION: STATUS OF AN ANIMAL

; FILE REFERENCE: DI-9-1

; CURRENT APPLICATION NUMBER: US/10/670,695

; CURRENT FILING DATE: 2003-09-25

; PRIOR APPLICATION NUMBER: 09/521,738

; PRIOR FILING DATE: 2000-03-09

; NUMBER OF SEQ ID NOS: 36

; SOFTWARE: Patent in Ver. 2.1

; SEQ ID NO 2

; LENGTH: 671

; TYPE: PRT

; ORGANISM: Feline calicivirus

US-10-670-695-2

Query Match

Best Local Similarity 9.5%; Score 275.5; DB 12; Length 671;

Matches 123; Conservative 70; Mismatches 174; Indels 185; Gaps 23;

Qy 67 SPNSPGEVLLNLELQEPNLYLAHARMYNGVAGFEVQVVLGNAGTAKIIFAAIPPP 126

Db 179 STSETQKILFKQSLGPLNLYLAHARMYNGVAGFEVQVVLGNAGTAKIIFAAIPPP 238

Qy 127 NFPIDNLSAQITMCPHVI DVROLEPVNPLMPDVNRNFFHYNQGSRLRLIAMLTYPL 186

Db 239 G--VDPVQSTSLMQYPHVLFDAQVEPVIFCPDLKSTLYHLMSDQDT--TSLVIMVYNDL 295

Qy 187 RANNSGDDVFTVSC--RVLTSPSPDFSNFLVPP-----TVES-----KT-----224

Db 296 -INPANDANSGCCIVVETKPGDPFKFHLKPPGSMTHGSIPLDLPKTSLSLWGNRY 354  
Qy 225 -----KPTTL-----PI-LTISEMNSRPPVPPIESLHTSP 253  
Db 355 WSDITDFVIRPVFOANRHDFNQETAGSTPRFRPISVTITEQNGAKLIGV-----A 408  
Qy 254 TENIVVQONG--RVTLDELGTTQLLPSQICAFRGVLTSTRSASDAQADTATPRLFNY 311  
Db 409 TDYIVPGIPDGWPDPTIPGEL-----IPAGDYAITNGTN-----443  
Qy 312 YHVHQLDNLNGTTPYPAEDI PGPLGTDPFRGVGVASORNLDSITTRAHEAKVDTTAGRF 371  
Db 444 -----DITATGYDTADIK--NNTNFRGMVYCGSLQRAWG-----DKISNTAFIT 488  
Qy 372 TPKGLSLEISTDSD--FDQOQTKFTPVGIGVDNEABFQWMSLPDYSQGFTH 422  
Db 489 TATL-----DGDNNKINPCNTIDQSKI VVQDNHVGKKAQTSDDTLALLGYTG----537  
Qy 423 NNNLAPAVAPNPFGEQLL-----FFRSQLPSSGSRNGVLDCLVPQEWVQH--FYQBS 473  
Db 538 -----IGEQAIGSDRDRVVRISTLPETGARG--NHPIFYKNS 573  
Qy 474 AP-----AQTOVALRVYVNP-----DTGKVLPEAKLHKLGMFT 506  
Db 574 IKLGVVIRSIDVFNSQLHTSRQLSLNHYLLPPDSFAVYRIIDNSGSMFIDIGSDGFSF 633  
Qy 507 IANNGDSPITVP 518  
Db 634 VGVSGFGKLEFP 645

## RESULT 5

US-10-209-507-2

; Sequence 2, Application US/10209507

; Publication No. US20030109033A1

; GENERAL INFORMATION:

; APPLICANT: Audonnet, et al.

; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT

; FILE REFERENCE: 454313-3151.2

; CURRENT APPLICATION NUMBER: US/10/209,507

; CURRENT FILING DATE: 2002-11-12

; PRIOR APPLICATION NUMBER: 09/617,594

; PRIOR FILING DATE: 2000-07-14

; PRIOR APPLICATION NUMBER: 60/193,332

; PRIOR FILING DATE: 2000-03-30

; PRIOR APPLICATION NUMBER: France 00 01761

; PRIOR FILING DATE: 2000-02-11

; PRIOR APPLICATION NUMBER: France 99 09421

; PRIOR FILING DATE: 1999-07-16

; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: Patent in version 3.0

; SEQ ID NO 2

; LENGTH: 669

; TYPE: PRT

; ORGANISM: Feline calicivirus

US-10-209-507-2

Query Match

Best Local Similarity 9.0%; Score 262; DB 14; Length 669;

Matches 110; Conservative 69; Mismatches 166; Indels 148; Gaps 20;

Qy 30 DPVAGAAATAAPLTGQONTIIDPMIMNFVQAPGGEFTV-----SPRNSPGEVLLNLELQEP 83

Db 146 EFSAQMATADATCKSVDSSEW--ESF-----FSFHTSVNWSSETQKILFKQSLGP 196

Qy 84 EINYLAHLARMYNGVAGFEVQVVLGNAGTAKIIFAAIPNFPIDNLSAQITMCPH 143

Db 197 LLNPYLEHLKLYVAVSGSDVVRFSISGSGVFGKLAALVWPPG--VDPVQSTSLMQYPH 254

Qy 144 VLVVROLEPVNPLMPDVNRNFFHYNQGSRLRLIAMLTYPL-----RANNSGDDVFTV 198

Db 255 VLFDAQVEPVIFSPDLRSTLYHLMSDQDT--TSLVIMVYNDLINPYANDNSSGCIVTV 313

Wed Jun 2 09:13:31 2004

QY 199 SCRVLTRSPDPSFNFVLP  
Db 314 E-----TKGPDKFKLLKPPGSMTHGSIPLSLPKSSLMWGNRYWSDITDFVIRFVF 369  
QY 230 -----PI-LTISEMNSRPPVPIESLHTSPTEINIVVOQNG-RV 266  
Db 370 QANRHDFNQETAGWSTFRFRPITITISENGSKLGTGV-----ADYIIVGIPDGPWD 423  
QY 267 TLDELGMGTQLLPSQICAFRGVLTSTRSRASDQADATPRLFNYWHVQDLNLTNGPYD 326  
Db 424 TTIGE-----ELTPA-----GDYSITNGSGNDIA-----TANAYD 453  
QY 327 PAEDIPGPGTDFRGKVGVSQORNLDSSTTRAHEAKVDVTTAGRTPKLSLEISTDSDD 386  
Db 454 SADVI---TNTNFRGMYTCGALQAWG-----DKKISSTA-FITTAKEGNTLKPSNT 503  
QY 387 FQONQPTKTPVIGIGVDNEAEFQQWLSLPDYSQGFTHNMNLAPAVAPNPFGEQLLFR--- 443  
Db 504 IDMTKAVQDTHVGRDVQTSDDTLAILGTG-----IGEQAIGSNRDS 547  
QY 444 ----SOLPSSGGR 452  
Db 548 VRIISMLPETGAR 560

RESULT 6  
US-10-369-493-2646  
; Sequence 2646, Application US/10369493  
; Publication No. US20030233675A1  
; GENERAL INFORMATION:  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Hinkle, Gregory J.  
; APPLICANT: Slater, Steven C.  
; APPLICANT: Goldman, Barry S.  
; APPLICANT: Chen, Xianfeng  
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF  
; TITLE OF INVENTION: PLANTS WITH IMPROVED PROPERTIES  
; FILE REFERENCE: 38-10(52052)B  
; CURRENT APPLICATION NUMBER: US/10/369,493  
; CURRENT FILING DATE: 2003-02-28  
; PRIOR APPLICATION NUMBER: US 60/360,039  
; PRIOR FILING DATE: 2002-02-21  
; NUMBER OF SEQ ID NOS: 47374  
; SEQ ID NO 2646  
; LENGTH: 1965  
; TYPE: PRT  
; ORGANISM: *Synechocystis* sp.  
US-10-369-493-2646

Query Match 4.4%; Score 126; DB 15; Length 1965;  
Best Local Similarity 20.9%; Pred. No. 0.061; Indels 186; Gaps 29;  
Matches 133; Conservative 76; Mismatches 241;  
QY 3 MASNDAAPSNDGAGIVPEINNEAMALDPVAGAAIAAPLTGQONIIDPWIMNPFVQAPGG 62  
Db 969 IATDDADVASAGITFGKLP-GND-----ACSFINPITGKVLIDD---PDFETQSSY 1016  
QY 63 EFTVSPR---NSFGEVLLNL-----ELGEPINPYLAHARMYNGYAGGFVQVVLGNA 113  
Db 1017 KFTVTASDGVNLATEQLVNLTIIDLEIAPTITSGDTAKAIAENSGANQVIYEIVATDDA 1076  
QY 114 FTAGKILFAIP-----PNPIDLSAAQITMCPHIVIVDVRQLE 152  
Db 1077 DVSAGITFGKLPKNDAGSFNITGKVLIDDPDFETQSSYKFTVTASDGVNLATEQL- 1135  
QY 153 PVNLPMEDVRN-----NFFHYNQGSDSKRLIAMLYPLRANNSG 192  
Db 1136 -VNLTIIDLEIAPTITSGDTAKAIAENSGANQIIYQATADDNADISAGVTGLKPGDDA 1194  
QY 193 DD--VFTVSCRVLTRSPDF-----SFNFLVPPPTVESKTFPTLITISEMNSRPPVPI 246  
Db 1195 DSFIINAITGQVLLNDPDFETQSSYKFTV--TASDGVNPVTEQVNVNLTIIDLEIAPTI 1252

QY 247 ESLHTSP--TEN-----IVVQCQNGRVTL-- 268  
Db 1253 TSGNTAKAIAENSGANQVIYQVATDDADISAEVTFGLGDDATSFIDSTTGKVTLLN 1312  
QY 269 --DGLMGTTQLLPSQICAFRGVLTSTRSRASDQADATPRLFNYWHVQDLNLT----- 320  
Db 1313 NPDFESQSSYK-----TWTASDGVNPATEKIVT-----LAINLDEIGPT 1353  
QY 321 --NGTPYDPAEDIPGP-----LCTPD---PRGKVGVSQORNLDSSTTRAHEAKVDVTTA 368  
Db 1354 ITSQDTATPIDENSGANQLIYQVIATDDADVASAGITFGI-----KPGNDA 1398  
QY 369 GRT--PKLSLEISTDSDDFQONQPTKTPVIGIGVDNEAEFQQWLSL-----DYSQQTH 422  
Db 1399 GSTIINPITGEVTL-IDDPDFETQSSYKFTVTASDGVNPATEKIVTLTINDLDELGNIT 1457  
QY 423 NMNLAPAVAPNPFGEQLLFRSQPSSGGR--SNGVLDCLVPOEWVQHFYQESAPAQTV 480  
Db 1458 SDNIATAINENSGANQIIY---QVTADDGDDISAGVTFGKPGDNANQFSID--PTTGQV 1512  
QY 481 ALVRYNPDPGKVLFEAKLHKLGFMTIANNQSPIT 516  
Db 1513 TLLD--NPD-----FETQ--SSYKFTVTASDGVNPAT 1540

RESULT 7  
US-10-172-502-17  
; Sequence 17, Application US/10172502  
; Publication No. US20030185833A1  
; GENERAL INFORMATION:  
; APPLICANT: FOSTER, Timothy et al.  
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES.  
; FILE REFERENCE: P07263US01/BAS  
; CURRENT APPLICATION NUMBER: US/10/172,502  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US 60/298,098  
; PRIOR FILING DATE: 2001-06-15  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 17  
; LENGTH: 522  
; TYPE: PRT  
; ORGANISM: *Staphylococcus epidermidis*  
US-10-172-502-17  
Query Match 4.2%; Score 121.5; DB 14; Length 522;  
Best Local Similarity 23.5%; Pred. No. 0.02;  
Matches 87; Conservative 39; Mismatches 131; Indels 113; Gaps 21;  
QY 219 TVESKTKPFTLPILITISEMNSRPPVPIESLHT-----SPT-----E 255  
Db 144 TSEARTSTNQISNLSTSTSNQSSPTSFANLTFPSRFTVLTMTAAPTTTTSSLSN 203  
QY 256 NIWVQCQ-----GRVLDGELMGTTQLLPSQICAFRGVLTSTRSRASDQADATPR 307  
Db 204 SVVVKDNFNEHNLNLSGSAITDPK-TGIATLTPDAY-SQKGAISLNLRLDSN-----R 254  
QY 308 LFNYWHVQDLNLTNGTPYD-----AEDI-----DPELGTDPFRGKVGVSQORNLDSST 357  
Db 255 SFRFIGKVNIGN-RYEGYSPDGVAGDGGIGFPAFGPLGQIGKEGAAGVIGLNN----- 308  
QY 358 RAHEAKVDVTTAGRTFKP-----LGSLEISTDSDDFQONQTKFTPVCI 400  
Db 309 -AGFKLDTYHTSTSPRSDAKAKADPRNVGGGAGFAG-VSTDRNGMATTESTAALKNV 366  
QY 401 -GVDNEAEFQQWLSLPDYSQ-----FTNMNLAPAVAPNPFGEQLLFRSQP 447  
Db 367 QPTDN--SFQDFVI-DYNGDTKVMVTYACQTFTRNLT---DWIKNSGG--TTFSLMWA 418  
QY 448 SSGRSNGVLDCLVPOEWVQHFYQESAPAQTVQALVRYNPDPGKVLFEAKLHKLGFMTI 507  
Db 419 STGGAKN-----LQOVQFGTFEYTESA-----VAKVRYVDANTGKDIIPPK-----TI 461

QY 508 ANNGDSPITV 517  
Db 462 AGEVDGTNI 471

RESULT 8

US-10-282-122A-71033  
; Sequence 71033, Application US/10282122A  
; Publication No. US20040029129A1

GENERAL INFORMATION:

; APPLICANT: Wang, Liangu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.

TITLE OF INVENTION: Identification of Essential Genes in Microorganisms

; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 71033

LENGTH: 1831

TYPE: PRT

ORGANISM: Staphylococcus epidermidis

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (1801)..(1801)

OTHER INFORMATION: X=any amino acid

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (1805)..(1805)

OTHER INFORMATION: X=any amino acid

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (1808)..(1808)

OTHER INFORMATION: X=any amino acid

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (1809)..(1809)

OTHER INFORMATION: X=any amino acid

FEATURE:

NAME/KEY: MISC FEATURE

LOCATION: (1825)..(1825)

; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (1828)..(1828)  
; OTHER INFORMATION: X=any amino acid  
US-10-282-122A-71033

Query Match 4.2%; Score 121.5; DB 12; Length 1831;  
Best Local Similarity 23.5%; Pred. No. 0.15;  
Matches 87; Conservative 39; Mismatches 131; Indels 113; Gaps 21;

QY 219 TVESKTKPTLPILTISEMSNSRFPVPIESLHT-----SPT-----E 255  
Db 236 TSEARTSTNQISNLSTSTSNOSPTSFLNLTFRFTVLNTMAAPTSTTTTSLTGN 295  
QY 256 NIVVOQN-----GRVTLDELMTGTTQLLPISOICAPRGVLTSTSTASQADATPR 307  
Db 296 SVVVKDNFNEHMLSGSATYDPK-TGIATLTTPDAY-SQKGAISLNLRLDSN-----R 346  
QY 308 LPNYYVHVOLDNLNGTPYDP-----AEDI-----PGLTDPFRGKVFVASORNLDSTT 357  
Db 347 SPRTGKVNLSN-RYEGYSPDGVAGDGGIGFAFSPGLQIGKEGAAVIGGLNN----- 400  
QY 358 RAHEAKVDTTAGRTPK-----LGSLEISTSDDDFDQONQTKTPVGI 400  
Db 401 -AFGFKLDTYHNTSTPRSDAKAKADPRNVGGGAGFAGF-VSTRNGMATTEETAALKV 458  
QY 401 -GVDNEAEFQOWSLPDYSGO-----FTNMNLAPAVAPNPFGEQLLFFRSQLP 447  
Db 459 QPTDN-SFQDFVI-DYNGDTKVTMTYAGQTFTRNLT---DWKNSGG--TTFSLMTA 510  
QY 448 SSGRSNGVLDCLVPQEWQHFYQBSAPAQTOVALRVYVNPDTGKVLPEAKHLKLGFM 507  
Db 511 STGGAKN-----LQVQFGTTEYTESA-----VAKRVYVDANTGKDIIPPK-----TI 553  
QY 508 ANNGDSPITV 517  
Db 554 AGEVDGTNI 563

RESULT 9

US-09-867-932A-2  
; Sequence 2, Application US/09867932A  
; Publication No. US20030049825A1

GENERAL INFORMATION:

; APPLICANT: Lin, Han-You  
; APPLICANT: Kuo, Tsun-Yung  
; APPLICANT: Huang, Hsiao-I  
; APPLICANT: Yang, Huey-Lang  
; TITLE OF INVENTION: NERVOUS NECROSIS VIRUS PROTEIN  
; FILE REFERENCE: 08919-023001  
; CURRENT APPLICATION NUMBER: US/09/867,932A  
; CURRENT FILING DATE: 2001-05-30  
; NUMBER OF SEQ ID NOS: 2  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 338  
; TYPE: PRT  
; ORGANISM: Nervous necrosis virus  
US-09-867-932A-2

Query Match 4.2%; Score 121; DB 10; Length 338;  
Best Local Similarity 24.6%; Pred. No. 0.011;  
Matches 82; Conservative 44; Mismatches 141; Indels 66; Gaps 17;

QY 8 AAPSDNGAAGLVP---EINNEAMALDPVAGAAIAAPLTG---QQNIIDPWIMNFWQA---P 60  
Db 10 AKPATTKAANPQRRRRANRRNRNRTDAPVSKASTVTGFGRCNTDNLHSLGMSRISQAVLP 69  
QY 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGVAG--GFEVQVLGNAFTAG 117  
Db 70 AGTGT-----DGYVVVDATIVPDLPLRLGHAARIFQRYAVETLEFEIQPMCFAN--TGG 121

143 --HVIIVDRQLEPV-NLMPDV---RNNFFHYNGSDSRLRL-----IA 180  
158 EDHVFNDSANWKPVTNLKAPQLKAGENHFSFDKVETAVIRVMVKADNRKGTISITEVQ 217  
181 MLYTFLRANNSGDDVFTVSCRVLTRPSPDFSNFLVPTVESKTKPFTLPILITISEMSNS 240  
218 IFAKQVAAAKQGTQRIQVDGKDLANFNPDLDYLL--ESVDGK-----VPAVTAS-VSN 269  
241 RFPVPFESLHTSPTENIVVQCONGRVTLTDLGELMGTQLLPSQICAFRGVLTSTRSRASDQ 300  
270 GLATVVPVSREGEPVRVIAKAEN-----GDILGEVRL--HFTKDKSLLSHKPVAAVKQ 320  
301 A-----DTATPRLFNYVHVLQNLNLTNGTYPD-PAEDIPGLTPTDPGRGVFGVA 348  
321 ARLLQVGQALELPTKVPVYFTKDGVTETKDLTVEWEVPAENLT-KAGQFTVRGVFLG-- 377  
349 SQRLNDSITRAHEAKVDTAGRTPKLGSLEISTDSDDFDQONQPTKFTPVGIGVDNEAEF 408  
378 --SNL-----VAEITVRVTDKLG--ETLSNPNVDENSNOAFASATNDIDKNSHD 423  
409 QQWNSLPDYSQOFTHN---MNLAPAVAPNFPFGQLLFFRSQLPSSGGRSNG-VLDCLVPOE 464  
424 RVDYLAND--GDHSENRRTWNSPTPSSNPEVSAGVIF-----RENGKIVERTVTOG 472  
465 WVQHFOESAPATOVALRVYVNPDTGKVLFEAKLHKLGMFTIANNNGDSPIITVPPN 520  
473 KVQFFADSGTDAPSKLVLERYVGPE-----FEVPTYISNYQ--AYDADHPFNPNPEN 521

RESULT 11  
US-09-769-787-2  
; Sequence 2, Application US/09769787  
; Publication No. US20030091577A1  
; GENERAL INFORMATION:  
; APPLICANT: Microbial Technics Limited  
; APPLICANT: Gilbert, Christophe FG  
; APPLICANT: Hansbro, Philip M  
; TITLE OF INVENTION: Proteins  
; FILE REFERENCE: PWC/P21129WO  
; CURRENT APPLICATION NUMBER: US/09/769,787  
; CURRENT FILING DATE: 2001-01-26  
; PRIOR APPLICATION NUMBER: GB 9816337.1  
; PRIOR FILING DATE: 1998-03-27  
; PRIOR APPLICATION NUMBER: US 60/125164  
; PRIOR FILING DATE: 1999-03-19  
; NUMBER OF SEQ ID NOS: 388  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 2233  
; TYPE: PRT  
; ORGANISM: Streptococcus pneumoniae  
US-09-769-787-2

Query Match 4.2%; Score 121; DB 10; Length 2233;  
Best Local Similarity 20.5%; Pred. No. 0.23;  
Matches 110; Conservative 75; Mismatches 233; Indels 118; Gaps 25;

31 PVAGAAIAAPLTGQQNIIDPWIMNFVQAPGGEFTVSPRNSPGEVLLNLELGPINPYLA 90  
1480 PLAFASDSNPSDPVSNVDKLIISYN--NQPANRWTNWRTNP-EASVGLFGD----- 1529

91 HLARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIPNFPIDNLSAAQITMCP----- 142  
1530 -----SGILSKESVDNLSVGFHEDHG-----VGVPKSYVIEYVGVGTPTAPKNPSFVGN 1579

143 --HVIIVDRQLEPV-NLMPDV---RNNFFHYNGSDSRLRL-----IA 180  
1580 EDHVFNDSANWKPVTNLKAPQLKAGENHFSFDKVETAVIRVMVKADNRKGTISITEVQ 1639  
181 MLYTFLRANNSGDDVFTVSCRVLTRPSPDFSNFLVPTVESKTKPFTLPILITISEMSNS 240  
1640 IFAKQVAAAKQGTQRIQVDGKDLANFNPDLDYLL--ESVDGK-----VPAVTAS-VSN 1691

118 KIIPAAIPNFPIDN---LSAAQITMCPHVIIVDRQLEPVNLMPDVNRNFFHYNGSDS 174  
122 GYVAGFLPD--PTDNDHTFGALQATR-CAVAKWWSERTVR--PQVTRTLLTSSGKEQ 175  
175 RLRLIAMLYTLPLRANNSGDDVFTVS--CRVLTRSPDFSNFLVPTVESKTKPFTLPIL 232  
176 RLTSFGRLLILCVGNNT--DVNVSVLCRWSVR-----LSVPSLETPEET-TAPIM 223  
233 TISEMSNRFPPIESLHTSPTENIVVQCONGRVTLTDLGELMGTQLLPSQICAFRGVLT 292  
224 TQGSYLN-----DSLTNDSKSLGSLTPLDIAPDGAVQLDRLLSIDYSLGTGDVDR 276  
293 STSRASDAQATATPRLFNYVHVLQNLNLTNGTYPD-PAEDIPGLTPTDPGRGVFGVA 324  
277 AV-----YHMLKKFAGNAGTP 292

US-09-765-272-86  
; Sequence 86, Application US/09765272  
; Patent No. US20020061545A1  
; GENERAL INFORMATION:  
; APPLICANT: Choi et. al.  
; TITLE OF INVENTION: Streptococcus pneumoniae Antigens and Vaccines  
; NUMBER OF SEQUENCES: 452  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Human Genome Sciences, Inc.  
; STREET: 9410 Key West Avenue  
; CITY: Rockville  
; STATE: Maryland  
; COUNTRY: USA  
; ZIP: 20850  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage  
; COMPUTER: HP Vectra 486/33  
; OPERATING SYSTEM: MSDOS version 6.2  
; SOFTWARE: ASCII Text  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/765,272  
; FILING DATE: 22-Jan-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/961,083  
; FILING DATE: <Unknown>  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Brookes, A. Anders  
; REGISTRATION NUMBER: 36,373  
; REFERENCE/DOCKET NUMBER: PB340P2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (301) 309-8504  
; TELEFAX: (301) 309-8512  
; INFORMATION FOR SEQ ID NO: 86:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 776 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 86:  
US-09-765-272-86

Query Match 4.2%; Score 121; DB 9; Length 776;  
Best Local Similarity 20.5%; Pred. No. 0.042;  
Matches 110; Conservative 75; Mismatches 233; Indels 118; Gaps 25;

31 PVAGAAIAAPLTGQQNIIDPWIMNFVQAPGGEFTVSPRNSPGEVLLNLELGPINPYLA 90  
58 PLAFASDSNPSDPVSNVDKLIISYN--NQPANRWTNWRTNP-EASVGLFGD----- 107  
91 HLARMYNGYAGGFEVQVVLGNAFTAGKIIFAAIPNFPIDNLSAAQITMCP----- 142  
108 -----SGILSKESVDNLSVGFHEDHG-----VGVPKSYVIEYVGVGTPTAPKNPSFVGN 157





Wed Jun 2 09:13:31 2004

us-09-926-799-6.rapb

Matches 96; Conservative 50; Mismatches 137; Indels 180; Gaps 23;  
QY 125 PRNF--PIDNLSAAITWCPHIVDVRLQLE-PVN-----LPMEDVRNPFHYNGS 172  
Db 149 PPSFQTPVNTVSSTNLVTPPAVVSQPKLQTPVTSGSLTATSVLPAPNTAT----- 199  
QY 173 DSRLLIAMLYTPLRANSGDDVFTVSCR-----VLTTP 206  
Db 200 -----VWATTQVP-----SGNQPTISLQPLVILHVPVAVSSQPLLQSHPGTLVTNPQ 249  
QY 207 SPDFSG-NFLVPPTVESKTK-PFTLPILTISMSNSRFPVIESLHTSPTEN-----I 257  
Db 250 SGNVEFISVQSPTVSGLTGKPNVSLPSPNPKPNVPSVPSIQRNPTASAPLGTTL 309  
QY 258 VVQC-----QNGRVTL-----DGLMGTTQLLPQ-----ICAFRGVLTSTRASDQ 300  
Db 310 AVQAVPTAHSIVQATRTSLTVGPGSLYSPSTRNGPIQMKIPISAF-----STSSAAEQ 363  
QY 301 ADTATPRLFNYYWHVQLDNLNGTTPYDPAEDIPLGLTDPFRGKVFVGASQRNLDSSTTRAH 360  
Db 364 NSNTTPIEN-----QTNKI-----IDASVSKKAADSTSQCG 395  
QY 361 EAK-----VDT-----AGRTPKLSLEISTDSDDFDQNPTK-----FTPVG 399  
Db 396 KATGSDSSGVIDLTMDDDESGASQDPKLNHTFVSTMSQSPVSRPLQPIQAPPLQPSG 455  
QY 400 IGVNDEAEFQWLSLPDYSGQFTHNMNLAFAV-----APNFPGEQLLFFRSQLPSSGGR 452  
Db 456 VPTSGSPQTHLLP--TAPTTVNVTHRPVTVQVTRLPVPRAPANHVVY--TTLP----- 507  
QY 453 SNGVLDCLVPQEWQHFYQBSAPAQ-----QVALRYVNP 488  
Db 508 -----APPAQAPLRGTVMQAPAVRQVNP 530

Search completed: June 1, 2004, 14:04:37  
Job time : 36.0315 secs

RESULT 15  
US-10-172-502-18  
; Sequence 18, Application US/10172502  
; Publication No. US20030185833A1  
; GENERAL INFORMATION:  
; APPLICANT: FOSTER, Timothy et al.  
; TITLE OF INVENTION: CROSS-REACTIVE MONOCLONAL AND POLYCLONAL ANTIBODIES . .  
; FILE REFERENCE: P07263U01/BAS  
; CURRENT APPLICATION NUMBER: US/10/172,502  
; CURRENT FILING DATE: 2002-06-17  
; PRIOR APPLICATION NUMBER: US 60/298,098  
; PRIOR FILING DATE: 2001-06-15  
; NUMBER OF SEQ ID NOS: 29  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 18  
; LENGTH: 485  
; TYPE: PRT  
; ORGANISM: Staphylococcus epidermidis  
US-10-172-502-18  
Query Match 3.9%; Score 111.5; DB 14; Length 485;  
Best Local Similarity 22.6%; Pred. No. 0.16;  
Matches 89; Conservative 50; Mismatches 169; Indels 85; Gaps 19;  
QY 169 NQGSRLRLIAMLYTPLRANSGDDVFTVSCRVLTRPSPDPSPFNFLVPTVESKTKPFT 228  
Db 85 NTSSSDTKSVASTSSTEQPTNSTQSTASNTSQSTPS--SVNLNKTSTTSTAP-- 141  
QY 229 LPILITISEMSNRFPVIESLHTSPTENIV-----QCONGRVTLTDLGE--LMGTQ 277  
Db 142 VKLRTESRLAMSTFASAATT--TAVTANTITVKNLKOYMTTSGNATYDQSTGIVLTQ 199  
QY 278 LLPSQICAFRGVLTSTRASDAQDADTAPRLFNYYWHVQLDNL-----NGTPYDPAEDIP 332  
Db 200 DAYSQ-----KGATLTGTRIDSNKS-----FHFGSKVNLGNKYGHNKGNGDGIGFAFSP 248  
QY 333 GPLGTDPFRGKVFVA-----SQRLDSTTRAHAEKVDIT-----AGRF-----T 372

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 46.3893 Seconds  
(without alignments)  
3289.030 Million cell updates/sec

Title: US-09-926-799-7  
Perfect score: 2845  
Sequence: 1 MKMASNDTPSNDGAAGLVP.....VNQFYSLAPMTGNGRRRVQ 540

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2845	100.0	540	AAB49706	Small rou
2	2269.5	79.8	535	AAB49707	Small rou
3	2018	70.9	548	AAB49705	Small rou
4	2001	70.3	550	AAB49709	Small rou
5	1979	69.6	548	AAB49708	Small rou
6	1933	67.9	542	AAB49704	Small rou
7	1872.5	65.8	539	AAB49704	Small rou
8	1650.5	58.0	541	AAB49710	Small rou
9	1233	43.3	530	AAB50972	Small rou
10	1233	43.3	530	ADC72176	Norwalk v
11	1232	43.3	530	AAB49701	Small rou
12	1225	43.1	545	AAB49700	Small rou
13	1183	41.6	530	AAR57091	Small rou
14	1170.5	41.1	544	AAB49703	Small rou
15	1169.5	41.1	546	AAB49702	Small rou
16	323.5	11.4	579	AAB08143	RHDV caps
17	296.5	10.4	668	AAB67462	Amino aci
18	292.5	10.3	547	AAM50108	Feline ca
19	292.5	10.3	668	AAR10686	Feline ca
20	292.5	10.3	668	AAB04304	Feline ca
21	292.5	10.3	671	AAM50107	Feline ca
22	281	9.9	622	AAB47045	Feline ca
23	281	9.9	623	AAB47044	Feline ca
24	277	9.7	623	AAB47043	Feline ca
25	273	9.6	669	AAB67461	Amino aci

26	186	6.5	40	5	AAU91274	Aau91274 Norwalk v
27	167	5.9	40	5	AAU91273	Aau91273 Norwalk v
28	139	4.9	934	1	AAP20016	Aap20016 Sequence
29	138	4.9	1147	5	ABF76724	Abf76724 Foot and
30	131.5	4.6	338	7	ADC06674	Adc06674 Nervous n
31	123.5	4.3	366	5	ABF05228	Abf05228 Redepotte
32	123	4.3	2209	1	ADP20037	Adp20037 Sequence
33	114.5	4.0	794	7	ADC03508	Adc03508 Ralstonia
34	111	3.9	6304	6	ABU09236	Abu09236 Human neu
35	110.5	3.9	1455	4	AAM79120	Aam79120 Human pro
36	109.5	3.8	689	6	ABR43116	AbR43116 Human IMX
37	109.5	3.8	756	7	ADC31441	Adc31441 Human nov
38	109.5	3.8	2701	5	ABP74125	Abp74125 Human TRI
39	109	3.8	257	6	ABU48168	Abu48168 Protein e
40	109	3.8	992	4	ABB60861	Abb60861 Drosophil
41	108.5	3.8	1125	3	AAB22934	Aab22934 Mouse mic
42	108.5	3.8	1125	3	AAV79637	Aav79637 Microtubu
43	108.5	3.8	1125	5	ABG94497	Abg94497 Protease
44	108.5	3.8	1125	7	ADC18464	Adc18464 Microtubu
45	108.5	3.8	1521	4	ABB52754	Abb52754 Escherich

ALIGNMENTS

RESULT 1  
AAB49706  
ID AAB49706 standard; protein; 540 AA.  
XX  
AC AAB49706;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 7.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX  
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
(DENK-) DENKA SEIKEN KK.  
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
WPI: 2001-080848/09.  
DR N-PSDB; AAF29147.  
PT Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.  
XX Claim 1; Page 54-57; 84pp; Japanese.  
PS This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks  
SQ Sequence 540 AA;  
Query Match 100.0%; Score 2845; DB 4; Length 540;  
Best Local Similarity 100.0%; Pred. No. 6.9e-256;



```
PD 28-DEC-2000.
XX
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST. INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI: 2001-080848/09.
XX N-PSDB; AAF29146.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX
XX Claim 1; Page 52-54; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 548 AA;
XX
XX Query Match 70.9%; Score 2018; DB 4; Length 548;
XX Best Local Similarity 68.3%; Pred. No. 9e-179;
XX Matches 377; Conservative 71; Mismatches 88; Indels 16; Gaps 4;
XX
XX QY 1 MKWASNDATPSNDGAGLVPESSNEMALEPVVGASLAAPVTGQTNIIDPWIRTFVQAP 60
XX DB 1 MKWASNDAAAPSNDGAGLVPEINNEAMALDPVAGAAIAPLFGQONIIDPWIRTFVQAP 60
XX
XX QY 61 NGEFTVSPRNSPGEILVNLELGPENLYLAHLARMYNGYAGGMEVQVVLGNAGNAGTAKII 120
XX DB 61 GGEFTVSPRNSPGEVLNLELGPENLYLAHLARMYNGYAGGFEVQVVLGNAGNAGTAKII 120
XX
XX QY 121 FAAPPPYFPVENLSQITMFPVHVIDVRLTLEPVLLPMPDVRSTLTFHFNQKDEPKMRLVA 180
XX DB 121 FAAPPPYFPIDNLSAAQITMCPHVIVDVRQLEPVLLPMPDVRNFFHYNQGSDSLRLITA 180
XX
XX QY 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDFTYLVPPVTESKTKPFTLPVLTGLGELS 240
XX DB 181 MLYTPLRANNSGDDVFTVSCRILTRPSDFSNFLVPPTVESKTKPFTLPVLTGLGELS 240
XX
XX QY 241 RPPLSIDEMVTSFNESIVVQPNQGRVTLTGELGTTQACNICISIRGKVTQV--PSEQ 298
XX DB 241 REPVIESLHTSPTENIVVQCNQGRVTLTGELGTTQACNICISIRGKVTQV--PSEQ 300
XX
XX QY 299 -----HWNLEITNLNGTQDPTDDVPAPLGVDPFAGEVGVLSQRNRSNAN 348
XX DB 301 ADATATPRLPNYWHVQDLNLTGTPYDPAEDIPGLGTDPGRKVGVSQRN---LDSTT 357
XX
XX QY 349 RAHDAAVATYSKYTPKGLVOIGTWNVDNENQTKFTPIGLNEVANCHREONTLPY 408
XX DB 358 RAHEAKVDTTAGRTFKPLGSLISIDSDDFQCNQPKTFPVGIG-VDNAAEFQOQSLPDY 416
XX
XX QY 409 SGALTNMNLAPAVAPLPGERLLPFRSVYVPLKGGFGNPAIDCSVPQEWQHFYQESAPS 468
XX DB 417 SQGFTHMNLAPAVAPNPFGEQLLFRSQLPSSGSGNSGVLDCLVPQEWQHFYQESAPA 476
XX
XX QY 469 LGDVALRVYNPDTGRVLFEAKLHKGGLTVTSSTGTPVVVPANGYFKFDSWVNFYSIA 528
XX DB 477 QTOVALRVYNPDTGRVLFEAKLHKLGFMTIANNGSDSPITVPNGYFRPESWVNFYTLA 536
XX
XX QY 529 PMGTGNRRRVQ 540
XX DB 537 PMGTGNRRRIQ 548
```

## RESULT 4

```
AAB49709
XX
XX AAB49709 standard; protein; 550 AA.
XX
XX AC AAB49709;
XX
XX DT 04-APR-2001 (first entry)
XX
XX DE Small round structured virus protein SEQ ID 10.
XX
XX KW Small round structured virus; SRSV; food poisoning.
XX
XX OS Small round structured virus.
XX
XX PN WO200079280-A1.
XX
XX PD 28-DEC-2000.
XX
XX PF 22-JUN-2000; 2000WO-JP004095.
XX
XX PR 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST. INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI: 2001-080848/09.
XX N-PSDB; AAF29150.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX
XX Claim 1; Page 62-64; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 550 AA;
XX
XX Query Match 70.3%; Score 2001; DB 4; Length 550;
XX Best Local Similarity 67.9%; Pred. No. 3.5e-177;
XX Matches 377; Conservative 72; Mismatches 86; Indels 20; Gaps 4;
XX
XX QY 1 MKWASNDATPSNDGAGLVPESSNEMALEPVVGASLAAPVTGQTNIIDPWIRTFVQAP 60
XX DB 1 MKWASNDAAAPSNDGAGLVPEANDEVMALPEPVVGASIAAPVVGQONIIDPWIRTFVQAP 60
XX
XX QY 61 NGEFTVSPRNSPGEILVNLELGPENLYLAHLARMYNGYAGGMEVQVVLGNAGNAGTAKII 120
XX DB 61 QGEFTVSPRNSPGEMLNLELGPENLYLAHLARMYNGYAGGQVQVVLGNAGNAGTAKII 120
XX
XX QY 121 FAAPPPYFPVENLSQITMFPVHVIDVRLTLEPVLLPMPDVRSTLTFHFNQKDEPKMRLVA 180
XX DB 121 FAAPPPYFPVENISAAQITMCPHVIVDVRQLEPVLLPDIERNRPFHYNQENTPRMLVA 180
XX
XX QY 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDFTYLVPPVTESKTKPFTLPVLTGLGELS 240
XX DB 181 MLYTPLRAN-SEGEDVFTVSCRILTRPAPDFEFTLVPPVTESKTKPFTLPVLTGLGELS 239
XX
XX QY 241 RPPLSIDEMVTSFNESIVVQPNQGRVTLTGELGTTQACNICISIRGKVTQV----- 295
XX DB 240 REPAIDMLYTDPNESIVVQPNQGRCTLDGTLGGTTLQVPTQICAFRGTLLISQTAARAAS 299
XX
XX QY 296 -----SEQHMNLEITNLNGTQDPTDDVPAPLGVDPFAGEVGVLSQRN-----RGSN 345
```

61 CGEFTVSPRNSPGFVLLNLELGPINFLAHLARMYNGYAGGFEVQVVLVLAFTAGKII 120  
 121 FFAVPPYFVNLSPQITMEPHVIIDVRLTEPVLPMPOVRSFLTFHFNQKDEKMLVA 180  
 121 FFAIPNPFIIDNLAAQITMCPHVIVRQLFVNLPMVVRNFFHYNQGSRLRLIA 180  
 181 MLYTLRNGSGDDVFTVSCRLTRPSPEFDTLVLPPTVESKTPPTLVLTGLSLSNS 240  
 181 MLYTLRANNSGDDVFTVSCRLTRPSPEFDTLVLPPTVESKTPPTLVLTGLSLSNS 240  
 241 RPLSIDEMVTSPNESIVVQONGRVLTDGELLGTTLQACNICSIIRKGVQGV--PSEQ 298  
 241 RFPVPTESLHTSPENIVVQONGRVLTDGELLGTTLQACNICSIIRKGVQGV--PSEQ 300  
 299 -----HMMLEITNLNGTQDDPDPAPLGVDPDFAGFVGVLSQNRGESNPAN 348  
 301 ADIATPRLFNYYHVQLDNLNGTQDDPDPAPLGVDPDFAGFVGVLSQNRGESNPAN 357  
 349 RAHDVAVATYSDKYTPKLGVLQVIGTWNTNDVENQKTPETIGLNEVANGHRPEQWTLPRY 408  
 358 RAHEAKVDITAGRTFKLSLEISTESSDFDQDQPTFRTPVGIG--VDNEADPQQWLSLPDY 416  
 409 SGALTLLNMLAPAVAPLFPGERLLPFRSYVPLKGGFGNPAIDCSVPQEWVQHYOESAPS 468  
 417 SQOFTHMLNMLAPAVAPLFPGERLLPFRSYVPLKGGFGNPAIDCSVPQEWVQHYOESAPS 476  
 469 LGDVALVRYVNPDTGRVLFPEAKLHKGGLTVSTSTGTPVVPANGYKFDSDVWNOFYSIA 528  
 477 QTQVALVRYVNPDTGRVLFPEAKLHKGGLTVSTSTGTPVVPANGYKFDSDVWNOFYSIA 536  
 529 PMGTGNGR 536  
 537 PMGTGNGR 544

RESULT 5  
 AAB49708  
 ID AAB49708 standard; protein; 542 AA.  
 XX  
 AC AAB49708;  
 XX  
 DT 04-APR-2001 (first entry)  
 XX  
 DE Small round structured virus protein SEQ ID 9.  
 XX  
 KW Small round structured virus; SRSV; food poisoning.  
 XX  
 OS Small round structured virus.  
 XX  
 PN WO200079280-A1.  
 XX  
 PD 28-DEC-2000.  
 XX  
 PF 22-JUN-2000; 2000WO-JP0040095.  
 XX  
 PR 22-JUN-1999; 99JP-00175928.  
 XX  
 PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 XX  
 PA (DENK-) DENKA SEIKEN KK.  
 XX  
 PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 XX  
 DR WPI; 2001-080848/09.  
 XX  
 DR N-PSDB; AAF29149.  
 XX  
 PT Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.  
 XX  
 PS Claim 1; Page 59-61; 84pp; Japanese.  
 XX  
 CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies

300 TDSPOARHHLHVQKXLDGTQYDPTDDIIPAVLGAIDFKGTVEGVASQRDVSCQBOQH 359  
 346 PANRAHDVAVATYSDKYTPKLGVLQVIGTWNTNDVENQKTPETIGLNEVANGHRFEQWTL 405  
 360 YATRAHEAHIDTDPKYAPKLGTLIKSGDDFNTNQPIRFTPVGMGD---NNRWQWEL 415  
 406 PRYSALTLNMLAPAVAPLFPGERLLPFRSYVPLKGGFGNPAIDCSVPQEWVQHYOES 465  
 416 PYSGLTLNMLAPAVAPLFPGERLLPFRSYVPLKGGFGNPAIDCSVPQEWVQHYOES 475  
 466 APSLGDVALVRYVNPDTGRVLFPEAKLHKGGLTVSTSTGTPVVPANGYKFDSDVWNOFY 525  
 476 APSQSAVALVRYVNPDTGRVLFPEAKLHKGGLTVSTSTGTPVVPANGYKFDSDVWNOFY 535  
 526 SLAPMTGNGRRRVQ 540  
 536 TLAPMGSGQGRRAQ 550

AAU91272  
 ID AAU91272 standard; protein; 548 AA.  
 XX  
 AC AAU91272;  
 XX  
 DT 18-JUN-2002 (first entry)  
 XX  
 DE Norwalk virus associated polynucleotide #1.  
 XX  
 KW Norwalk virus; monoclonal antibody; geno group I; geno group II;  
 KW immunological detection; food; viral infection.  
 XX  
 OS Norwalk virus.  
 XX  
 PN JP2002020399-A.  
 XX  
 PD 23-JAN-2002.  
 XX  
 PF 10-JUL-2000; 2000JP-00208151.  
 XX  
 PR 10-JUL-2000; 2000JP-00208151.  
 XX  
 PA (OSAP ) OSAKA PREFECTURE.  
 PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.  
 PA (IATR ) IATRON LAB INC.  
 XX  
 DR WPI; 2002-287412/33.  
 XX  
 PT A monoclonal antibody useful in the immunological detection and diagnosis  
 PT of Norwalk virus infection.  
 XX  
 XX Disclosure; Page 12-13; 24pp; Japanese.  
 XX  
 CC The invention describes a monoclonal antibody recognising Norwalk virus,  
 CC a capsid protein of Norwalk virus, or a common antigen epitope on the  
 CC capsid protein molecule of geno group I and geno group II. The antibody  
 CC is useful for immunological detection and quantitative analysis of  
 CC Norwalk virus in foods and the serum of infected patients. This sequence  
 CC represents a Norwalk virus associated protein described in the invention  
 XX  
 XX Sequence 548 AA;  
 SQ  
 Query Match 59.6%; Score 1979; DB 5; Length 548;  
 Best Local Similarity 67.7%; Pred. No. 3.9e-175;  
 Matches 371; Conservative 71; Mismatches 90; Indels 16; Gaps 4;  
 QY 1 MKMASNDATPSNDGAAGLVPEESNEAMALEPVPVVGASLAAPVTGQTIIDPWIRTFVQAP 60  
 DB 1 MKMASRAAFPSNDGAAGLVPEESNEAMALDPVAGAAIAAPLTGQQIIDPWIMNPFVQAP 60  
 QY 61 NGFTVSPRNSPGFVLLNLELGPINFLAHLARMYNGYAGGFEVQVVLVLAFTAGKII 120

CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX  
XX Sequence 542 AA;  
SQ  
Query Match 67.9%; Score 1933; DB 4; Length 542;  
Best Local Similarity 67.0%; Pred. No. 7.4e-171;  
Matches 366; Conservative 76; Mismatches 94; Indels 10; Gaps 6;  
QY 1 MKWASNDATPSNDGAGLVPESSNEAMALEPVVGASLAAPVTGQTNIIIDPWRTNPFVQAP 60  
DB 1 MKWASNDATPSNDGAGLVPESSNEAMALEPVVGASLAAPVTGQTNIIIDPWRTNPFVQAP 60  
QY 61 NGFEVSPRNSPGEILVNLELGPDLNPNYLALHARMYNGVAGMEVQVLMAGNAFTAGKII 120  
DB 61 NGFEVSPRNSPGEILVNLELGPDLNPNYLALHARMYNGVAGMEVQVLMAGNAFTAGKII 120  
QY 121 FAAPVPYFVENLSPSQITMFPFHVIIIDVRLPVLPMDFVSTLTFHFQKDEPKMRLVA 180  
DB 121 FAAPVPYFVENLSPSQITMFPFHVIIIDVRLPVLPMDFVSTLTFHFQKDEPKMRLVA 180  
QY 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDFTYLVPPVTSKTKPFTLPVLTLGELSNS 240  
DB 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDFTYLVPPVTSKTKPFTLPVLTLGELSNS 240  
QY 241 RPLSIDEMVTSPNESIVVQPNQGRVTLDELGLTTLQACNICISIRGVKTCQVPSQHM 300  
DB 241 RPLSIDEMVTSPNESIVVQPNQGRVTLDELGLTTLQACNICISIRGVKTCQVPSQHM 300  
QY 301 WN---LEITNLNGTQDPTDDVPAPLGVDPDFAGEVGVLSQRNGES-NPANRAHDVA 474  
DB 301 WN---LEITNLNGTQDPTDDVPAPLGVDPDFAGEVGVLSQRNGES-NPANRAHDVA 474  
QY 475 VRYVNPDTGRVLFKALHKGGLTYSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTCN 534  
DB 475 VRYVNPDTGRVLFKALHKGGLTYSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTCN 534  
QY 535 GRRRVQ 540  
DB 537 GRRRIQ 542  
RESULT 7  
ID AAB49704  
XX AAB49704 standard; protein; 539 AA.  
AC AAB49704;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 5.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
FF 22-JUN-2000; 2000WO-JP004095.  
XX

PR 22-JUN-1999; 99JP-00175928.  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
XX  
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX  
DR WPI; 2001-080848/09.  
DR N-PSDB; AAF29145.  
XX  
PT Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX  
PS Claim 1; Page 50-52; 84pp; Japanese.  
XX  
CC This invention relates to a kit for the detection and typing of small  
CC round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX  
SQ Sequence 539 AA;  
Query Match 65.8%; Score 1872.5; DB 4; Length 539;  
Best Local Similarity 64.5%; Pred. No. 3.2e-165;  
Matches 349; Conservative 80; Mismatches 105; Indels 7; Gaps 4;  
QY 1 MKWASNDATPSNDGAGLVPESSNEAMALEPVVGASLAAPVTGQTNIIIDPWRTNPFVQAP 60  
DB 1 MKWASNDATPSNDGAGLVPESSNEAMALEPVVGASLAAPVTGQTNIIIDPWRTNPFVQAP 60  
QY 61 NGFEVSPRNSPGEILVNLELGPDLNPNYLALHARMYNGVAGMEVQVLMAGNAFTAGKII 120  
DB 61 NGFEVSPRNSPGEILVNLELGPDLNPNYLALHARMYNGVAGMEVQVLMAGNAFTAGKII 120  
QY 121 FAAPVPYFVENLSPSQITMFPFHVIIIDVRLPVLPMDFVSTLTFHFQKDEPKMRLVA 180  
DB 121 FAAPVPYFVENLSPSQITMFPFHVIIIDVRLPVLPMDFVSTLTFHFQKDEPKMRLVA 180  
QY 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDFTYLVPPVTSKTKPFTLPVLTLGELSNS 240  
DB 181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDFTYLVPPVTSKTKPFTLPVLTLGELSNS 240  
QY 241 RPLSIDEMVTSPNESIVVQPNQGRVTLDELGLTTLQACNICISIRGVKTCQVPSQHM 300  
DB 241 RPLSIDEMVTSPNESIVVQPNQGRVTLDELGLTTLQACNICISIRGVKTCQVPSQHM 300  
QY 301 WNLEITNLNGTQDPTDDVPAPLGVDPDFAGEVGVLSQRNGESNPANRAHDVA 474  
DB 301 WNLEITNLNGTQDPTDDVPAPLGVDPDFAGEVGVLSQRNGESNPANRAHDVA 474  
QY 475 VRYVNPDTGRVLFKALHKGGLTYSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTCN 534  
DB 475 VRYVNPDTGRVLFKALHKGGLTYSSTSTGVPVVPANGYFKFDSWVNOFYSLAPMGTCN 534  
QY 535 GRRRVQ 540  
DB 537 GRRRIQ 542  
RESULT 8  
ID AAB49710  
XX AAB49710

DB	359	GGDHLCPQISSEIYLTSPNLTCTNPQPLPQSLGRTILIRSDNGHCHDMVGTSPITPT	418
QY	380	ENQPTKFTPIGLNEVANGHRFEQWTLPRYSGALTINNNLAPAVAPLPPGERLLFFRSVP	439
DB	419	WPQWRCSRSGNSCCSGHRY-----PVPVWNRVTWIV-----LS	454
QY	440	LKGGFGNPAIDCSVPQ---EWQHFYQBSAPSLGDAVALRVYVNPDTGRVLFEAKLHKGF	496
DB	455	HKSGFSTST--RKLPLNLRW-----PLIRFNPDTGRVLFEARLHKQF	497
QY	497	LTVSSTSTGPVVVPVANGYFKFDSWYNQFYSYSLAPMTGNGRRVQ	540
DB	498	ITVAHTGDNPIVMPNGYFRFEAWNQFYSLAPVGTGKRRVQ	541
RESULT 9			
AA50972	1	AA50972 standard; protein; 530 AA.	
XX	AA50972;		
AC	16-OCT-2003 (revised)		
DT	25-MAR-2003 (revised)		
DT	05-OCT-1994 (first entry)		
XX	Norwalk virus strain 8FIIa protein (encoded by ORF2).		
DE	Norwalk virus; pathogen; acute gastroenteritis; food poisoning;		
XX	seafood contamination; diagnostic assay; calcivirus; small round virus.		
KW	Norwalk virus; (strain 8FIIa).		
XX	WO9405700-A2.		
PN	17-MAR-1994.		
XX	07-SEP-1993; 93WO-US008447.		
PF	07-SEP-1992; 92US-00941365.		
XX	(BAYU) BAYLOR COLLEGE MEDICINE.		
PR	Matson DO, Estes MK, Jiang X, Graham DY;		
XX	WPI: 1994-101125/12.		
XX	N-PSDB; AAQ56826.		
DR	DNA from Norwalk and related viruses - used for preparing prods. for use		
PT	in diagnostic assays, detection and vaccines for Norwalk and related		
PT	viruses.		
XX	Claim 14; Page 68-70; 156pp; English.		
PS	The Norwalk virus was isolated from stool samples from adult volunteers		
XX	infected with safety tested Norwalk virus strain 8FIIa. The coding		
CC	sequence is useful for the design of probes for use in diagnostic assays		
CC	for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct		
CC	PN field.) (Updated on 16-OCT-2003 to standardise OS field)		
XX	Sequence 530 AA;		
QY	Query Match 43.3%; Score 1233; DB 2; Length 530;		
DB	Best Local Similarity 47.6%; Pred. No. 1.3e-105;		
DB	Matches 263; Conservative 81; Mismatches 163; Indels 46; Gaps 11;		
QY	1 MKMASNDATPSNDGAAG---LVPESN-NEAMALEPVGASLAAPVTGQNIIDPWIRTNF	56	
DB	1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWLINF	60	
QY	57 VQAPNGEFTVSPNSGCEILVNLELGPENLPYLALARMYNGYAGGMEVQVMLAGNAFTA	116	
DB	61 VQAPQGEFTTSPNNTPGDVLFDLSLGPLHLPNPFLLHLSQMYNGVGNVRVIMLAGNAFTA	120	
AA50972	1	AA50972 standard; protein; 530 AA.	
XX	AA50972;		
AC	16-OCT-2003 (revised)		
DT	25-MAR-2003 (revised)		
DT	05-OCT-1994 (first entry)		
XX	Norwalk virus strain 8FIIa protein (encoded by ORF2).		
DE	Norwalk virus; pathogen; acute gastroenteritis; food poisoning;		
XX	seafood contamination; diagnostic assay; calcivirus; small round virus.		
KW	Norwalk virus; (strain 8FIIa).		
XX	WO9405700-A2.		
PN	17-MAR-1994.		
XX	07-SEP-1993; 93WO-US008447.		
PF	07-SEP-1992; 92US-00941365.		
XX	(BAYU) BAYLOR COLLEGE MEDICINE.		
PR	Matson DO, Estes MK, Jiang X, Graham DY;		
XX	WPI: 1994-101125/12.		
XX	N-PSDB; AAQ56826.		
DR	DNA from Norwalk and related viruses - used for preparing prods. for use		
PT	in diagnostic assays, detection and vaccines for Norwalk and related		
PT	viruses.		
XX	Claim 14; Page 68-70; 156pp; English.		
PS	The Norwalk virus was isolated from stool samples from adult volunteers		
XX	infected with safety tested Norwalk virus strain 8FIIa. The coding		
CC	sequence is useful for the design of probes for use in diagnostic assays		
CC	for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct		
CC	PN field.) (Updated on 16-OCT-2003 to standardise OS field)		
XX	Sequence 530 AA;		
QY	Query Match 43.3%; Score 1233; DB 2; Length 530;		
DB	Best Local Similarity 47.6%; Pred. No. 1.3e-105;		
DB	Matches 263; Conservative 81; Mismatches 163; Indels 46; Gaps 11;		
QY	1 MKMASNDATPSNDGAAG---LVPESN-NEAMALEPVGASLAAPVTGQNIIDPWIRTNF	56	
DB	1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWLINF	60	
QY	57 VQAPNGEFTVSPNSGCEILVNLELGPENLPYLALARMYNGYAGGMEVQVMLAGNAFTA	116	
DB	61 VQAPQGEFTTSPNNTPGDVLFDLSLGPLHLPNPFLLHLSQMYNGVGNVRVIMLAGNAFTA	120	
AA50972	1	AA50972 standard; protein; 530 AA.	
XX	AA50972;		
AC	16-OCT-2003 (revised)		
DT	25-MAR-2003 (revised)		
DT	05-OCT-1994 (first entry)		
XX	Norwalk virus strain 8FIIa protein (encoded by ORF2).		
DE	Norwalk virus; pathogen; acute gastroenteritis; food poisoning;		
XX	seafood contamination; diagnostic assay; calcivirus; small round virus.		
KW	Norwalk virus; (strain 8FIIa).		
XX	WO9405700-A2.		
PN	17-MAR-1994.		
XX	07-SEP-1993; 93WO-US008447.		
PF	07-SEP-1992; 92US-00941365.		
XX	(BAYU) BAYLOR COLLEGE MEDICINE.		
PR	Matson DO, Estes MK, Jiang X, Graham DY;		
XX	WPI: 1994-101125/12.		
XX	N-PSDB; AAQ56826.		
DR	DNA from Norwalk and related viruses - used for preparing prods. for use		
PT	in diagnostic assays, detection and vaccines for Norwalk and related		
PT	viruses.		
XX	Claim 14; Page 68-70; 156pp; English.		
PS	The Norwalk virus was isolated from stool samples from adult volunteers		
XX	infected with safety tested Norwalk virus strain 8FIIa. The coding		
CC	sequence is useful for the design of probes for use in diagnostic assays		
CC	for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct		
CC	PN field.) (Updated on 16-OCT-2003 to standardise OS field)		
XX	Sequence 541 AA;		
QY	Query Match 58.0%; Score 1650.5; DB 4; Length 541;		
DB	Best Local Similarity 56.5%; Pred. No. 1.6e-144;		
DB	Matches 330; Conservative 65; Mismatches 102; Indels 87; Gaps 8;		
QY	1 MKMASNDATPSNDGAAGLVPSNNEAMALEPVGASLAAPVTGQNIIDPWIRTNFVQAP	60	
DB	1 MKMASNDAAAPSDGAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRTNFVQAP	60	
QY	61 NGEFTVSPNSGCEILVNLELGPENLPYLALARMYNGYAGGMEVQVMLAGNAFTAGKII	120	
DB	61 AGEFTVSPNSGCEILLDLELGPDLNPLYLALARMYNGHAGGMEVQIIVLAGNAFTAGKII	120	
QY	121 FFAVPPYPVFNLSQITFMFHVHIDVTLEPVLPMDDYRSTLFHFNQDEPKMLVA	180	
DB	121 FFAIPGPFYENLSPSQITMCPHVHIDVRQLEPFLPMPDIWNNEFFHYNQNGDKLRLVA	180	
QY	181 MLYTPLRNSGDDVFTVSCRILTRPSPEFDTYLVPTVESKTKPFTLPVLTIGELNS	240	
DB	181 MLYTPURANNSGDDVFTVSCRVLTKPSDFETFLVPTVESKTKQFALPILKISEMTNS	240	
QY	241 RFLPSIDEMVTSNIVVQPNQGRVTLTGELLTQLOACNICIRGKV---TQGVSE	297	
DB	241 RFPVPDVMYATARNQVQPNQGRVTLTGELLTGPLLVNICKFGEVIANKGDVRS-	299	
QY	298 QHWNLEITNLNGTQFDPTDDVPAPLGPDPFAGEVGLVSQRNKGESNPANRAHDAVVAT	357	
DB	300 -YRMDMEITNTGDTPTDPTEDTGPICGDPDFQGLFEGVASQRNKGEPATRAHEAIINT	358	
QY	358 YSDKYTPKLGIVQI-----GTWNTNDV	379	





Accession	Source	Length	Score	DB	Indels	Mismatches	Conservative	Best Local Similarity	Query Match
28-DEC-2000.		530	43.3%	Score 1232;	DB 4;	Length 530;			
22-JUN-2000; 2000WO-JP004095.		530	47.3%	Pred. No. 1.6e-105;					
22-JUN-1999; 99JP-00175928.		530	81;	Mismatches 157;	Indels 56;	Gaps 12;			
(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.		530							
(DENK-) DENKA SEIKEN KK.		530							
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;		530							
WPI; 2001-080848/09.		530							
N-PSDB; AAF29142.		530							
Kit for the detection and typing of small round-structured virus (SRSV)		530							
strains for investigation of food poisoning outbreaks, contains		530							
antibodies.		530							
Claim 1; Page 42-45; 84pp; Japanese.		530							
This invention relates to a kit for the detection and typing of small		530							
round structured virus (SRSV) strains. The kit contains antibodies		530							
directed against peptides represented in sequences AAB49700 - AAB49710,		530							
which are each SRV strain specific. Polynucleotide sequences AAF20141 -		530							
AAF20151 represent cDNA encoding the strain specific proteins. The kit is		530							
used for detecting and typing strains of SRV in order to prevent the		530							
spread of infection and to examine the epidemiology of outbreaks		530							
Sequence 530 AA;		530							
Query Match		530							
Best Local Similarity		530							
Mismatches 264; Conservative		530							
1 MKMASNDATPSNCGAAG---LVPESN-NEAMALEPVGASLAAPVTGQTNIDPWRNF 56		530							
1 MMASKAOTSSVCGASGAGQLVPEVNASDPLANDPVAGSSTAVATAGQVNPIDPWIINF 60		530							
57 VQAPNGEFTVSPRNSPGEILVNLELGPENLPIVLAHARMYNGVAGMEVQVMLAGNAFTA 116		530							
61 VQAPQGEFTISPNPTGGLVFDLSGLPHLNPFLHLSQMYNGVGNVRVIMLAGNAFTA 120		530							
117 GKIIFAAVPPYFPVENLSPSQITMFPHVIDVRLPEVLLPMPDVSTLPHFNQKDEPRM 176		530							
121 GKIIIVSCIPPGFSGHNLITIAQLFPFHVIADVRVLEPFLVFNHNRNQTM 180		530							
177 RLVMALYPLRSNGSGDDVFTVSCRILTRPSPEFDFTYLVPPVTSKTPFTLPVLTLGE 236		530							
181 RLVCMLYPLRTGGGTGDSFVAGRVMTCPSPDNFLVPPVTEQTRPFTLPNPLSS 240		530							
237 LNSRFPISIDBMTVSPNESIVVQNGRVTLGDLGTTQLQACNCSIRGKVTQVPS 296		530							
241 LNSRAPLPISGMIGISPNVQSVQNGRCVTLGRLVTPVLSHVAKIRGTNGTV-- 298		530							
297 EQHMWNLITNLNGTQFDPTDDVPAPLGVDPFAG-----EVEGLVSQRNRESNPANR 349		530							
299 -----INULEDGTGTFHPFEG-PAPIGFDLGGCDWHINMTQGHSSQTQ----- 342		530							
350 AHDVAVATYSDKYTKPLGLVQ---IGTWNVDNENQPTKFTPIGNEVA-----NGHREE 401		530							
343 -YD--VDTPPTFVPHLGSIQANGIGSGNYIGV-----LSWVSPSPHSGSQVD 398		530							
402 QWTLPRYSGALTNLNNLAPVAPLPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWQHF 461		530							
389 LMKIPNYGSSITEATHLAPSVTPPGFGEVLVFFMSKIPGAGYSLP---CLLPQEYISHL 445		530							
462 YQESAPSLGVALRVYVNDTGRVLPFAKLHKGGLTV--SSTSTGPPVVPANGYFKFDS 519		530							
446 ASQAQPTVGEAALLHVVDPTGRTLGEFRAYPDGELTICVPNGASSGPQQLPVGVPVS 505		530							
520 WYNQFVSLAPMGTCNGER 537		530							
506 WVSFRYQKPVGTASSAR 523		530							

Db 300 GNGY-----NLTELGSPYHAFES-PAPIGFPLD-GECDWHM-----EASPTTQFTQGD 347  
 QY 355 V-----ATYSKYTKPLGLVQIGTWTNNDVENOPTKFTPIG-LNEVANGHR--FEQWTL 406  
 Db 348 VIKQINVKQESAFAPHLGTIQAD--GLSDSVNTNMIAKLGWVSPVSDGHRGDVDPWIP 405  
 QY 407 RYSGALTLMNMLAPAVAPLPPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWYQYBSA 466  
 Db 406 RYGSTLTEAAQLAPPIYPGFGGEAIVFFMSDPPIAHGTNGLSVPCCTIPOEFVTFVNEQA 465  
 QY 467 PSLGDVALRVYNPDGTGRVLFPAKLHKGGLTV--SSTSTGPPVVVPANGYFKFDSWVQPF 524  
 Db 466 PRGGAALLHYLDDPDTHRLNLGFKLYPEGFMTCVPSNCGTGPOTLPINGVFVFSWVSRF 525  
 QY 525 YSLAPMGTCN-----GRRR 538  
 Db 526 YQLKPVGTAGPACRLGIR 544

## RESULT 13

AAR57091

ID AAR57091 standard; protein; 530 AA.

XX AC AAR57091;

XX DT 27-AUG-2003 (revised)

XX DT 25-MAR-2003 (revised)

XX DT 05-OCT-1994 (first entry)

XX DE Small round virus SRSV/KY/89 capsid protein.

XX KW pathogen; acute gastroenteritis; food poisoning; seafood contamination;

XX KW diagnostic assay; human calcivirus; small round virus; SRSV; KY89;

XX KW Norwalk virus; capsid protein.

XX OS Small round structured virus.

XX PN WO9405700-A2.

XX PD 17-MAR-1994.

XX PF 07-SEP-1993; 93WO-US008447.

XX PR 07-SEP-1992; 92US-00941365.

XX PA (BAYU) BAYLOR COLLEGE MEDICINE.

XX PI Matson DO, Estes MK, Jiang X, Graham DY;

XX DR WPI; 1994-101125/12.

XX DR N-PSDB; AAQ56832.

XX PT DNA from Norwalk and related viruses - used for preparing prods. for use

XX PT in diagnostic assays, detection and vaccines for Norwalk and related

XX PS viruses.

XX PS Example 7; Fig 13a; 156pp; English.

XX CC The known sequence for Norwalk virus was used to obtain the sequence of

XX CC other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool

XX CC from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide

XX CC cDNA sequence includes part of the polymerase region and the capsid

XX CC region of the genome; the deduced amino acid sequences are AAR57092 and

XX CC AAR57091, respectively. Expression of fragments and derivs. of Norwalk-

XX CC related viruses permits development of diagnostic assays to detect

XX CC antibodies, antigens, viral genetic material or antivirals. (Updated on

XX CC 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS

XX CC field.)

XX SQ Sequence 530 AA;

Query Match 41.6%; Score 1183; DB 2; Length 530;

Best Local Similarity 46.3%; Pred. No. 6e-101;  
 Matches 259; Conservative 84; Mismatches 158; Indels 58; Gaps 14;  
 QY 1 MKMASNDATPSNDGAAG--LVPESN-NEAMALEPVGASLAAPVTGOTNIIDPWIRTNF 56  
 Db 1 MMASKDATSVVDGASASVQLVPEVNASDPLAMDPAVGSSTAVATAGQVNPIDPHIINF 60  
 QY 57 VOAPNGEFTVSPRNSPGEILNVLELGPBEPNLYLAHARMYNGYAGGMEVQVNLAGNAFTA 116  
 Db 61 VOAPQGEFTISPNTPGDVLFDLSLGPBEPNLYLAHARMYNGYAGGMEVQVNLAGNAFTA 120  
 QY 117 GKIIIPAAVPVPVENLSQITMPPHVIIDVRLPVLMPDVRSTLPHNQKDEPKM 176  
 Db 121 GKIIIVSCIPPGFGSQLTIAQATLPPHVIADVRTLDPIEVPLEDVRNVLPHNDRNQOTM 180  
 QY 177 RLVAMLYTPLRNSGSGDDVFTVSCRILTRPSPEFDFTLVLPVPTVSKTKPFTLPVLTIGE 236  
 Db 181 RLVCMLYTPLTSGGTGDSFVAVGRVMTCPSPDFNFLVLPPTVQKTRPFTLPNLPSS 240  
 QY 237 LNSRFPLSIDEMVTSNPESIVVQPNQGRVTLDSLLGTTQLQACNICIRKVTGQVPS 296  
 Db 241 LNSRAPLPISGMGISPDNVQSVQFQNGRCTLDGLRVGTTPVSLSHVAKIRGTSNGTV-- 298  
 QY 297 EOHMNLITNLNGTQFDPDTPDVPAPLGVDPFAG-----EVFGVLQNRNCGESNPANR 349  
 Db 299 -----INLTLDGTPFHPFEG-PAPIGFPDLGGCDWHINMTQFGHSSQTQ----- 342  
 QY 350 AHDAVVATYSKYTKPLGLVQ--IGTWTNNDVENOPTKFTPIGLNEVA-----NGHRE 401  
 Db 343 -YD--VDTPTDTSVPHLGSIQANGIGSGNYIGV-----LSWSPSPHSGSQVD 388  
 QY 402 QMTLPYSGALTLNMLNLAAPV-APLPPGERLLFFRSYVPLKGGFGNPAIDCSVPQEWYQ 460  
 Db 389 LMKIPNYGSSITEATHLAPSVVSPGF-GEVLVFPMSKIP--GPGGDSLPCLLPQGYISH 444  
 QY 461 FYQESAPSLGDVALRVYNPDGTGRVLFPAKLHKGGLTV--SSTSTGPPVVVPANGYFKF 518  
 Db 445 LASEQAPTVGEGPLLYVDPTDRNLGFKAYPDGFLTCVPNGASSGPOQLPINGVFVFPV 504  
 QY 519 SWVNOFYSLAPMGTCNGRR 537  
 Db 505 SWVSRFYQLKPVGTASTAR 523

## RESULT 14

AAB49703

ID AAB49703 standard; protein; 544 AA.

XX AC AAB49703;

XX DT 04-APR-2001 (first entry)

XX DE Small round structured virus protein SEQ ID 4.

XX KW Small round structured virus; SRSV; food poisoning.

XX OS Small round structured virus.

XX PN WO200079280-A1.

XX PD 28-DEC-2000.

XX PF 22-JUN-2000; 2000WO-JP004095.

XX PR 22-JUN-1999; 99JP-00175928.

XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

XX PA (DENK-) DENKA SEIKEN KK.

XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX DR WPI; 2001-080848/09.

XX DR N-PSDB; AAF29144.

```

XX PT Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX PS
XX PS Claim 1; Page 47-49; 84pp; Japanese.
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX CC
XX CC Sequence 544 AA;
XX SQ
XX Query Match 41.1%; Score 1170.5; DB 4; Length 544;
XX Best Local Similarity 45.6%; Pred. No. 9.2e-100; Indels 53; Gaps 16;
XX Matches 258; Conservative 85; Mismatches 170;
XX
XX QY 1 MKMASNDATPSNDGAAG--LVPESEN-EAMALEPVVGASLAAPVTGQNIIDPWIRTNF 56
XX DB 1 MMWASKDAPSDAGATGAGLVPEVNTADPLPMEFVAGPTTAVATAGQVNNIDPWVNNP 60
XX
XX QY 57 VQAPNGEFTVSPRNSPGEILVNLELGPENLYLAHARMYNGVAGMEVQVMLAGNAFTA 116
XX DB 61 VQAPQGEFTISPNTPGDVLFDLQGLPHLNPFLSHLSQMYNGVGNMRVRVVLGNAFTA 120
XX
XX QY 117 GKIIFAAPVPPFVFNLSPSQITMFPFHVIIIDVRLPVLPMDDVRSITLHFHKNQDEPKM 176
XX DB 121 GKVIICCVPPFGQSRITLSIAQATLFPHVIAADVRLDPVEVLEDRVNLVH-NNDFQPTM 179
XX
XX QY 177 RLVMALYTPLRNSGS--GDDVFTVSCRILTRPSPDFDYLVPPVTVESKTPFLPVLT 234
XX DB 180 RLICMLYTPLRNGASGTDTSFVAGRVLTCPGDFNFLVLPVTEQKTRPPTVFNIP 239
XX
XX QY 235 GELSNRPPLSIDMWTSFNESIVVQPNQGRVTLDELLGTTLQACNICIRGKVTQGV 294
XX DB 240 KYLSNRIPNPIEGMSLSPQTNVQFQNGRCTIDGQPLGTTTPSVSQCCKFRGIT--- 296
XX
XX QY 295 PSEOHMNLLETNLNGTQFDPTDDVPAPLGVDPDPAGEVFGVLSQRNRGESNPANRAHDAV 354
XX DB 297 -SGQRVNL--TELDGSPF-MAFAPAPAGFPDLGSCDWHIEMSKIPINSSTQNNPIVTS 352
XX
XX QY 355 VATYSDKYTPKGLIV-----QIGT--WNT--NDVENQTKFTPIGLNEVANGHR 399
XX DB 353 VKPNSQQFVPHLSITLSDENSSGGDYIGTIQWTSPPSDSGGANTNF----- 399
XX
XX QY 400 FEQWTLPRYSQALTLNMNLAPAVAPLPFG--ERLLFFRSYVPLKGGFGNP-AIDCSVPQE 456
XX DB 400 ---WKIPDYGSLAASQALAPVY--PGFNEVIVYFWASIPGPNQSGSPNLVPCLLPQE 454
XX
XX QY 457 WYQHFYQBSAPSLGDVALRVYNDPTGRVLPEAKLHKGGLTV--SSTSIGVVPVANGY 514
XX DB 455 YITHFISEQAPIQGEAALLHYVDPDTRNLGFEKLYPGGYLTCTVFNSSSTGPOQLPLDGV 514
XX
XX QY 515 FKFDWVWQVSLAPMGT-GNGRRV 539
XX DB 515 FVFASWVSRYQLKPVGTAGPARGRL 540
XX
XX RESULT 15
XX AAB49702
XX ID AAB49702 standard; protein; 546 AA.
XX AC
XX AC AAB49702;
XX AC
XX DT 04-APR-2001 (first entry)
XX DE
XX DE Small round structured virus protein SEQ ID 3.
XX KW Small round structured virus; SRSV; food poisoning.

```

```

XX OS Small round structured virus.
XX PN WO200079280-A1.
XX PD 28-DEC-2000.
XX PF 22-JUN-2000; 2000WO-JP004095.
XX PR 22-JUN-1999; 99JP-00175928.
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX PA (DENK-) DENKA SEIKEN KK.
XX PI Takeda N, Natori K, Miyamura T, Kanata K, Sato T, Sato S;
XX DR WPI; 2001-080848/09.
XX DR N-PSDB; AAF29143.
XX PT Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX PS Claim 1; Page 45-47; 84pp; Japanese.
XX CC This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX CC
XX CC Sequence 546 AA;
XX SQ
XX Query Match 41.1%; Score 1169.5; DB 4; Length 546;
XX Best Local Similarity 44.9%; Pred. No. 1.1e-99;
XX Matches 254; Conservative 87; Mismatches 178; Indels 47; Gaps 13;
XX
XX QY 1 MKMASNDATPSNDGAAG--LVPESEN-EAMALEPVVGASLAAPVTGQNIIDPWIRTNF 56
XX DB 1 MMWASKDAPQADGASGAGQLVPEVNTADPLPMEFVAGPTTAVATAGQVNNIDPWVNNP 60
XX
XX QY 57 VQAPNGEFTVSPRNSPGEILVNLELGPENLYLAHARMYNGVAGMEVQVMLAGNAFTA 116
XX DB 61 VQSPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNMRVILLAGNAFSA 120
XX
XX QY 117 GKIIFAAPVPPFVFNLSPSQITMFPFHVIIIDVRLPVLPMDDVRSITLHFHKNQDEPKM 176
XX DB 121 GKIIICCVPPFGTSSSLTIAQATLFPHVIAADVRLPVEPLEDRVNLVHTND-NQPTM 179
XX
XX QY 177 RLVMALYTPLRNSGS--DVFTVSCRILTRPSPDFDYLVPPVTVESKTPFLPVLT 234
XX DB 180 RLVMCLYTPLRNGSGSDSFVAGRVLTAPSSDFSLFLVPPPTIEQKTRPPTVFNIP 239
XX
XX QY 235 GELSNRPPLSIDMWTSFNESIVVQPNQGRVTLDELLGTTLQACNICIRGKVTQGV 294
XX DB 240 QTLNSRFPFLSQGLLSEDSAQVVFQNGRCLIDGQLLGTTPATSGQLFRVGRKI---- 295
XX
XX QY 295 PSEOHMNLLETNLNGTQFDPTDDVPAPLGVDPDPAGEVFGVLSQRNRGESNPANRAHDAV 354
XX DB 296 --NQGARTLNLTEVDGKPFMAFDS-PAPVGFDPFGKCDWHMRISKTPNNTSSGDPMSVS 352
XX
XX QY 355 VATYSDKYTPKGLVQIGTWNINDVENQPT-----KFTPIGLNEVANGHRFEQ 402
XX DB 353 VQTNVQGFVPHLSIQF-----DEVFNHPTGDIYGTIEWISQSPSTPGTD-----INL 400
XX
XX QY 403 WTLFRYSQALTLNMNLAPAVAPLPFG--ERLLFFRSYVPLKGGFGNP-AIDCSVPQEWVHF 461
XX DB 401 WEIPDYGSSLSQAANLAPPVFPFGFGEALVYFSAFFGPNRNSAPNDVPCLLPQEVITF 460
XX
XX QY 462 YQBSAPSLGDVALRVYNDPTGRVLPEAKLHKGGLTV--SSTSIGVVPVANGYPKFDS 519

```

Db 461 VSEQPTMGDAALLHYVDDTNRNLGEFKLYPGGYLTCVPNGVGAGPQOLPLNGVFLFVS 520

Qy 520 WYNQFYSLAPMGTGN-----GRRRV 539

Db 521 WVSRYQLKPVGTASTARSRLGVRI 546

Search completed: June 1, 2004, 13:45:56  
Job time : 48.3893 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.8658 Seconds  
(without alignments)  
2166.837 Million cell updates/sec

Title: US-09-926-799-7  
Perfect score: 2845  
Sequence: 1 MKMASNDATPSNDGAAGLVP.....VNQFYS LAPMTGNGRRRVQ 540

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/prodata/2/iaa/5A COMB.pep.\*
- 2: /cgn2\_6/prodata/2/iaa/5B COMB.pep.\*
- 3: /cgn2\_6/prodata/2/iaa/6A COMB.pep.\*
- 4: /cgn2\_6/prodata/2/iaa/6B COMB.pep.\*
- 5: /cgn2\_6/prodata/2/iaa/PCTUS COMB.pep.\*
- 6: /cgn2\_6/prodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	43.3	530	4	US-08-486-049-3
2	296.5	10.4	688	4	US-09-617-594A-4
3	294.5	10.4	626	4	US-09-590-020-7
4	281	9.9	622	4	US-09-590-020-6
5	281	9.9	623	4	US-09-590-020-4
6	277	9.7	623	4	US-09-590-020-2
7	273	9.6	669	4	US-09-617-594A-2
8	137	4.8	2318	3	US-09-091-219-24
9	137	4.8	2318	4	US-09-660-541-24
10	111.5	3.9	2206	1	US-07-852-260-2
11	111.5	3.9	2206	2	US-08-461-503-2
12	111.5	3.9	2206	3	US-08-465-250-2
13	108.5	3.8	1125	4	US-09-513-783A-152
14	108.5	3.8	1610	4	US-09-513-783A-22
15	105.5	3.7	776	4	US-09-489-039A-13081
16	104.5	3.7	627	4	US-09-345-473E-46
17	104.5	3.7	968	4	US-09-228-986-76
18	104.5	3.7	2232	3	US-09-091-219-25
19	104.5	3.7	2232	4	US-09-660-541-25
20	104.5	3.7	2247	3	US-09-091-219-2
21	104.5	3.7	2247	4	US-09-660-541-2
22	103	3.6	519	4	US-09-489-039A-7785
23	101.5	3.6	878	1	US-08-237-919-2
24	101.5	3.6	878	4	US-08-732-429-2
25	101.5	3.6	878	4	US-09-798-267-2
26	101.5	3.6	878	4	US-09-798-267-3
27	101.5	3.6	878	5	PCT-US95-05518-2

28	101.5	3.6	2736	4	US-09-252-991A-30227	Sequence 30227, A
29	100	3.5	853	4	US-09-489-039A-11009	Sequence 11009, A
30	100	3.5	1091	6	5516630-2	Patent No. 5516630
31	100	3.5	2227	3	US-08-475-886-2	Sequence 2, Appl
32	100	3.5	2227	3	US-08-475-886-4	Sequence 4, Appl
33	100	3.5	2227	3	US-08-475-886-6	Sequence 6, Appl
34	100	3.5	2227	3	US-08-397-232-2	Sequence 2, Appl
35	100	3.5	2227	3	US-08-397-232-4	Sequence 4, Appl
36	100	3.5	2227	3	US-09-171-387-2	Sequence 2, Appl
37	100	3.5	2227	4	US-09-653-499-2	Sequence 4, Appl
38	100	3.5	2227	4	US-09-653-499-4	Sequence 6, Appl
39	100	3.5	2227	4	US-10-135-988-2	Sequence 2, Appl
40	100	3.5	2227	4	US-10-135-988-4	Sequence 4, Appl
41	100	3.5	2227	4	US-10-135-988-6	Sequence 6, Appl
42	100	3.5	1043	3	US-08-928-361B-30	Sequence 30, Appl
43	99.5	3.5	1721	3	US-08-700-651-5	Sequence 5, Appl
44	99.5	3.5	1721	3	US-08-928-361B-6	Sequence 6, Appl
45	99.5	3.5	1721	3	US-08-928-361B-6	Sequence 6, Appl

ALIGNMENTS

RESULT 1  
US-08-486-049-3  
; Sequence 3, Application US/08486049  
; Patent No. 6572862  
; GENERAL INFORMATION:  
; APPLICANT: Estes, Mary K  
; APPLICANT: Jiang, Xi  
; APPLICANT: Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE:  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-486-049-3

Query Match 43.3%; Score 1233; DB 4; Length 530;  
Best Local Similarity 47.6%; Pred. No. 2.5e-111;  
Matches 263; Conservative 81; Mismatches 163; Indels 46; Gaps 11;  
QY 1 MKMASNDATPSNDGAAG---LYPESN-NEAMALEPVVGASLAAPVTGQNIIDPWRTNF 56  
DB 1 MMWASKDATSSVDGASGAGQLVFEVNASDPLAMDVPVAGSSTAVATAGQVNPIDPWNNP 60

57 VOAPNGBTSPRNSGCEILVNLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTA 116  
 61 VOAPQGEFTISPNPTFGVDLSDLSQYNGWGNMVRIMLAGNAFTA 120  
 117 GKIIFAAVPPYPVENLSPSQITMPPHVIIDVRLTPVLLPMPDVRSTLFHFNOKDEPKMLV 179  
 121 GKIIIVSCIPPGSGHNLITQAATLPPHVIADVRLTDPIEVLDPVLEDRVNLPHNNDRNQTM 180  
 177 RLIVAMLYTLRNSGDDVFTVSCIRLTPSPPEFDTYLVPPPTVESKTKPFTLPVLITGE 236  
 181 RLIVCMLYTLRGTGGTGDSEFVAGRVMTCPSPDFELVLPVTEQKTRPFTLPNPLSS 240  
 237 LNSRFPPLSIDEMVTSNPSIIVVQPNQGRVTLDDGELLGTQLOACNICISIRGKVTQVPS 296  
 241 LNSRAPLPISMGISPDNVQVQFQNGRCTLDRGLVGTTPVSLSHVAKIRGTSNGTV-- 298  
 297 EOHMMNLEITNLNGTOFDTDDVPAPLGPVDFAG-----EVFGVLRSORNGESNPANR 349  
 299 -----INLTDLGTFFHPEG-PAPIGFPLDGGCDWHINMTQFGHSSQTQ----- 342  
 350 AHDVAVATYSKYTKPLGLVQ---IGTWNNDVENQPTKFTPIGLNEVANGHRFBQWTLF 406  
 343 -YD--VDTTPTDTPVPHLGSIQANGISGNYGV-----LSWISPPSHSGSQVDLMKIP 393  
 407 RYSGALTLNNLAPAVAPLPPGERLLFFRSYVPLKGGFGNPAIDCSVPOEWQHFVQESA 466  
 394 NYGSSITEATHLAPSVYVPGFGEVLVFFMSKMPGAYNLP---CLLPQEYISHLASEQA 450  
 467 PSIGDVALVRYNPDTGRVLPEAKLHKGGLTV--SSTSTGPRVVPANGYKFDGSMVNOF 524  
 451 PTVGEAALLHYVDPDTRNLNLFKAYPDGFLTCVPNGASSGPOQLFNGVFFVFSWVSFR 510  
 525 YSLAPMGTGNRR 537  
 511 YQLKPVGTASSAR 523

RESULT 2  
 US-09-617-594A-4  
 ; Sequence 4, Application US/09617594A  
 ; Patent No. 6541458  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Audonnet, et al.  
 ; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
 ; FILE REFERENCE: 454313-3151.1  
 ; CURRENT APPLICATION NUMBER: US/09/617,594A  
 ; CURRENT FILING DATE: 2000-07-14  
 ; PRIOR APPLICATION NUMBER: 60/193,332  
 ; PRIOR FILING DATE: 2000-03-30  
 ; PRIOR APPLICATION NUMBER: France 00 01761  
 ; PRIOR FILING DATE: 2000-02-11  
 ; PRIOR APPLICATION NUMBER: France 99 09421  
 ; PRIOR FILING DATE: 1999-07-16  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 668  
 ; TYPE: PRT  
 ; ORGANISM: Feline calicivirus  
 US-09-617-594A-4

Query Match 10.4%; Score 296.5; DB 4; Length 668;  
 Best Local Similarity 26.0%; Pred. No. 6.6e-20;  
 Matches 114; Conservative 55; Mismatches 135; Indels 135; Gaps 22;  
 11 SNDGAAGLVPSNNEAMALEPVGASLAAPVTGQTNIIIDPWIRTNFVOAPNGEFTV---- 66  
 125 ADDGSV-TTPEQGT-----LVGGVIAEPNAQMSAVAD--VATG--KSDVSEWEAFPSF 172  
 67 -----SPRNSGCEILVNLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTA 120  
 173 HTSVNWMSTETQKILFKQSLGPLNLPYLTHLAKLYVAVMSGSIEVRFSSISGSGVFGGKLA 232

121 FRAVPPYF-PVENLSPSQITMPPHVIIDVRLTPVLLPMPDVRSTLFHFNOKDEPKMLV 179  
 233 AIIVPPGIDPVGOSTMLQ---YPHVLFDFARQVEPVIFITPDRNSLYHL-MSDITDTTSLV 288  
 180 AMLYTTL-----RNSGDDVFTVSCIRLTPSPPEFDTYLVPPPTVESKTKPFTLPVLIT 233  
 289 IMIYNDLIIMPYANDSNSSG-CIVTVE---TKPGDPFKHLLKPPG-----SMLT 333  
 234 LGELNSRFPPLSIDEMVTSNPSIIVVQPNQGRVTLDDGELLGTQLOACNICISIRGKVTQ 293  
 334 HSGI-----PSDLI----- 342  
 294 VPEQHMMNLEITNLNGTOFDTDDVPAPLGPVDFAGEVGV-----LSQRNGESNPAN 348  
 343 -PKSSSLW---IGNRHS--DITDFVIKPF-----VFQANRHDFDFNQETAGWSTPRF 388  
 349 RAHDVAVATYSKYTKPLGLVQIGT-----WNTNDVENQPTKFTPIGLNEVANGH 398  
 389 R---PITITVSEKGGSKLG-IGVATDSIVPGIPDGMPTTI---PEKLTAPAGDYAITNGG 441  
 399 RFEQWTLPRYSGALTLNMM 417  
 442 NNDITTAADYDGASIIKQN 460

RESULT 3  
 US-09-590-020-7  
 ; Sequence 7, Application US/09590020  
 ; Patent No. 6355246  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Kruger, John M  
 ; APPLICANT: Maes, Roger K  
 ; APPLICANT: Vilnis, Aivars  
 ; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
 ; FILE REFERENCE: MSU 4.1-446  
 ; CURRENT APPLICATION NUMBER: US/09/590,020  
 ; CURRENT FILING DATE: 2000-06-08  
 ; PRIOR APPLICATION NUMBER: 60/138,484  
 ; PRIOR FILING DATE: 1999-06-10  
 ; NUMBER OF SEQ ID NOS: 9  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 7  
 ; LENGTH: 626  
 ; TYPE: PRT  
 ; ORGANISM: Feline calicivirus  
 US-09-590-020-7

Query Match 10.4%; Score 294.5; DB 4; Length 626;  
 Best Local Similarity 22.1%; Pred. No. 9.3e-20;  
 Matches 131; Conservative 84; Mismatches 220; Indels 157; Gaps 24;  
 11 SNDGAAGLVPSNNEAMALEPVGASLAAPVTGQTNIIIDPWIRTNFVOAPNGEFTVS 67  
 80 ADDGSI-TAPEQGTWVGGVIAEPSAQMSAADMATKGSVDSEWEAFSFTSVNW----S 134  
 68 PRNSGCEILVNLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAFTAIIFAAVPEY 127  
 135 TSETQKILFKQSLGPLNLPYLTHLAKLYVAVMSGSIEVRFSSISGSGVFGGKLAIVVPP- 193  
 128 FVENLSPSQITMPPHVIIDVRLTPVLLPMPDVRSTLFHFNOKDEPKMLVLYTPLR 187  
 194 -GVDPQVSTMLQYPHVLFDFARQVEPVIFCLPDLASTLYHL-MSDITDTTSLVIMVNDL- 250  
 188 SNGSGDDVFTVSC--RILTRPSPEFDTYLVPPPTVESKTKPFTLPVLITGELNSRFPPLS 245  
 251 INFYANDTSSGSGCIVVETKPGDPKDFHLLKPPG-----SMLTSGSVPSDLIPKS 300  
 246 -----IDEMVTSP-----NE-----SIVVQPNQNGR----- 265  
 301 SSLWIGNRHSWSDITDFIIRPFVQANRHDFDNQETAGWSTPRPRPISVITTEQNGAKLGI 360



Qy	266	-----VTLDGELLGTTQLQACH-----ICSTRGKGTQGV-----	294
Db	361	GVATYIVIEGIPDGWPDFTTIFGELIPAGDYAITNGTGDITATGYDTADIIKNNTNFRG	420
Qy	295	-----PSEQHMNLEITNLNGTQFOPTDDVPAPLGVPDFAGEVFGVLISQRNRGSESNAPNRA	350
Db	421	MYICGSLQRAWGDK-----KISNTAFITT-----ATLDGDNNNKINECNNTI	461
Qy	351	HDAVATYSDKYTPKGLGVQIGTWNNDVENOPTKFTPIGLNEVANGHREPO-----WTLP	406
Db	462	DQSKVIVFQDAHVGKKA-----QTSDDTLALLGYT--GIGQAIGSGDRDRVRIISTLP	512
Qy	407	RYSGALLTLNMNLAPAVAPLFPCEIRLLFPERSYVPLKGGFGNPADICSPQEWQHFQESA	466
Db	513	E-TGAR-----CGNHPIFYKSIKL--GYVIRSIDV-----FNSQIL	546
Qy	467	PSLGDVALVRYVNPDTGRVLFEAKLHGGFLTVSSSTGTPVVPVVPANGYFKFD	518
Db	547	HFSRQLSNLHYLLLPDPSFAVYRIIDNSGWSFPDIDIGDSGFSFVGVSQFGKLE	598

## RESULT 4

```

US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Villis, Aivars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590.020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-6

```

## RESULT 6

[illegible]

```

RESULT 8
US-09-091-219-24
; Sequence 24, Application US/09091219
; Patent No. 6171592
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/091,219
; CURRENT FILING DATE: 1998-10-05
; EARLIER APPLICATION NUMBER: PCT/AU96/00815
; EARLIER FILING DATE: 1996-12-18
; EARLIER APPLICATION NUMBER: AU PN7201
; EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-09-091-219-24

```

PRIOR APPLICATION NUMBER: 60/138,484  
PRIOR FILING DATE: 1999-06-10  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Patent in Ver. 2.1  
SEQ ID NO 2  
LENGTH: 623  
TYPE: PRT  
ORGANISM: Feline calicivirus  
US-09-590-020-2

Query Match 9.7%; Score 277; DB 4; Length 623;  
Best Local Similarity 24.8%; Pred. No. 4,7e-18;  
Matches 100; Conservative 53; Mismatches 142; Indels 108; Gaps 17;

QY 30 EPVVGASLAAPVYTGOTNIIDPW-IRNFVOAPNGEFTVSPRNSPGEIILVNLELGPENPY 88  
DB 100 EPSQAASAAADWATGKSDVSEWEAFPSFHTSVNW-----STSETQKILFKQSLGLLPNY 155  
QY 89 LAHLARMYNGYAGGMEVQVLMAGNAFTAGKIIFAAVPPYPVENLSPSOITMFPHVIIDV 148  
DB 156 LSHLAKLYVAVWSGVSEVRFPSISGSGVGGKLAIVVPP--GIEPQVSTMLQYPHVLFDA 213  
QY 149 RTLEPVLLPMPDVRSITLFHFNQKDEKMRVAMLYTPLRSNGSGDDVFTVSC--RILTRP 206  
DB 214 RQVEPVIFAIDRLSNLYHL-MSD TDTTSLVINVMYNDL-INPYANDTNSGGCIVTVKXP 271  
QY 207 SPSEDFTYLVPPVESKTKFTPLVLTGLGELSNRFPPLSIDEMVTSFNESIIVVQPNQGRV 266  
DB 272 GPDPKFHLLKPP-----GSMILT-----HGSV 292  
QY 267 TLGELLGTTQLQACNICSRGKVTGQVPSQHMWNLEITNLNGTQDPTDDVPAPLGPV 326  
DB 293 PSD-----LIPKSSSLW-----IGNRHS--DITDFIIRPF--- 321  
QY 327 DFAGEFQV-----LSQRNGESNPANRAHDVAVVATYDVKYTKLGLVQIGT----- 373  
DB 322 -----VFQANRHFDPNQTAGWSTPRFR---PITIIVSESNMKSGLG-IGVATDVIYVGP 372  
QY 374 --WNTNDVENOPTKFTPIGLNEVANGHRFEQWTLPRYSGALTL 414  
DB 373 DGMWPTTI---PEQLITPAGIYSITASNGTIVTTAAGYDAAETI 412

RESULT 7  
US-09-617-594A-2  
Sequence 2, Application US/09617594A  
Patent No. 6541458  
GENERAL INFORMATION:  
APPLICANT: Audomnet, et al.  
TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
FILE REFERENCE: 454313-3151.1  
CURRENT APPLICATION NUMBER: US/09/617,594A  
CURRENT FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/193,332  
PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: France 00 01761  
PRIOR FILING DATE: 2000-02-11  
PRIOR APPLICATION NUMBER: France 99 09421  
PRIOR FILING DATE: 1999-07-16  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: Patent in version 3.0  
SEQ ID NO 2  
LENGTH: 669  
TYPE: PRT  
ORGANISM: Feline calicivirus  
US-09-617-594A-2

Query Match 9.6%; Score 273; DB 4; Length 669;  
Best Local Similarity 22.4%; Pred. No. 1.3e-17;  
Matches 120; Conservative 74; Mismatches 188; Indels 154; Gaps 23;

QY 67 SPRNSPGEIILVNLELGPENPYLAHLARMYNGYAGGMEVQVLMAGNAFTAGKIIFAAVPP 126

Db 364 CHLLELPDTHKGVYGLTSDSYAMRNGWDVEVTAVGNQFNQGLLVAMVPPELYSTQKREL 423  
QY 136 SQITMPFHHVIDVRLPEVLLPMPDVRSTLFHFNQKDEPKMR-----LVAMLYTPPLRSNGS 191  
Db 424 YQLTLFPHQFINPRTNMTAHTVPEF-----GVNRYDQVKVHKPWTLVVMVVAPLTVNTE 478  
QY 192 GDDVTVSCRIILTRSPEDFTYLVPPPTVESKTKPFTPLVLTGELSISR--FPLSIDB- 248  
Db 479 G-----APQIKVYANIAPT-----NVHVGEPFSPKSGIFPVACSDG 514  
QY 249 ----MVTSPNESIVVQPNQGRV-----TLDCGELLGTTLOACNICSIRGKVTGOVP 295  
Db 515 YGGLVTTDPK---TADPVYGVKVFNPRLNQLGRF--TNLLDVAEACPTFLRPEGGVP 566

## RESULT 9

US-09-660-541-24  
; Sequence 24, Application US/09660541  
; Patent No. 6531136  
; GENERAL INFORMATION:  
; APPLICANT: STUDDERT, Michael J.  
; APPLICANT: CRABB, Brendan S.  
; APPLICANT: FENG, Li  
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS  
; FILE REFERENCE: 040268/0151  
; CURRENT APPLICATION NUMBER: US/09/660,541  
; CURRENT FILING DATE: 2000-09-12  
; PRIOR FILING DATE: EARLIER FILING DATE: 09/091,219  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201  
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 2318  
; TYPE: PRT  
; ORGANISM: Foot-and-mouth disease virus  
US-09-660-541-24

Query Match 4.8%; Score 137; DB 4; Length 2318;  
Best Local Similarity 21.6%; Pred. No. 0.0017;  
Matches 77; Conservative 40; Mismatches 130; Indels 110; Gaps 14;  
QY 11 SNDGAAGLVP-----ESNNE-----AMALEPVVGVASLAAPVTGQTNIIDPMTWTFVQA 59  
Db 248 SNEGSDTDTTSTHTTQNDWFSKLASSAFSLGALLADKKTETTLLEDRLIT-----T 303  
QY 60 NGEETVSPRSPG-----EILVN----- 78  
Db 304 RNHTTSTTQSSVGTGYGATAEDFVSGPNTSGLETRVVQAEPRFKTHLFDWVTSDFGR 363  
QY 79 ---LELGPENLPYLAHARMYNGYAGGMEVQVYMLAGNAFTAGKIIFAAVPPYFPVNLSP 135  
Db 364 CHLLELPDTHKGVYGLTSDSYAMRNGWDVEVTAVGNQFNQGLLVAMVPPELYSTQKREL 423  
QY 136 SQITMPFHHVIDVRLPEVLLPMPDVRSTLFHFNQKDEPKMR-----LVAMLYTPPLRSNGS 191  
Db 424 YQLTLFPHQFINPRTNMTAHTVPEF-----GVNRYDQVKVHKPWTLVVMVVAPLTVNTE 478  
QY 192 GDDVTVSCRIILTRSPEDFTYLVPPPTVESKTKPFTPLVLTGELSISR--FPLSIDB- 248  
Db 479 G-----APQIKVYANIAPT-----NVHVGEPFSPKSGIFPVACSDG 514  
QY 249 ----MVTSPNESIVVQPNQGRV-----TLDCGELLGTTLOACNICSIRGKVTGOVP 295  
Db 515 YGGLVTTDPK---TADPVYGVKVFNPRLNQLGRF--TNLLDVAEACPTFLRPEGGVP 566

## RESULT 10

US-07-852-260-2  
; Sequence 2, Application US/07852260  
; Patent No. 5525715  
; GENERAL INFORMATION:

; APPLICANT: Racaniello, Vincent  
; APPLICANT: Tatem, Joanne M.  
; APPLICANT: Weeks-Levy, Carolyn L.  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM  
; TITLE OF INVENTION: CDNA  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 30 Rockefeller Plaza  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/07/852,260  
; APPLICATION NUMBER: 07/852,260  
; FILING DATE: 19920619  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US  
; TELEPHONE: (212) 977-9550  
; TELEFAX: (212) 664-0525  
; TELEX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2206 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-07-852-260-2  
Query Match 3.9%; Score 111.5; DB 1; Length 2206;  
Best Local Similarity 17.9%; Pred. No. 0.47;  
Matches 107; Conservative 74; Mismatches 170; Indels 247; Gaps 28;  
QY 11 SNDGAAGLVP-----ESNNEAMALSPV-----VGASLAAPVT----- 42  
Db 270 TNSGATVLPVYNALALSDVMVKNWNGIALPLSLDPAQDSSVEIPITVTIAPMCSEFN 329  
QY 43 GQTNLIDP-----WIRTNFVQAPNG--EFTVSPR-NSPGEI----- 75  
Db 330 GLRNVTAPEKFOGLPVLNTPGNSQYLTSNHOQSCAIPFVDVTPPIDIFGEVKNMMLAEI 389  
QY 76 -----LVNLELGPENLPYLAH-----LARMYNG 98  
Db 390 DTMLPMLNLESTKRTMDMYRVTLSDSADLSQPICLSLSPADPRLSHTMLGEVLNYTH 449  
QY 99 YAGMEVQVMLAGNATAGKIIFAAVPPYFPVNLSPSQITMPFHHVIDVRLPEVLLPM 158  
Db 450 WAGSLKFTFLPCGSMATKILVAYAPGAQPT-SRKEAMLGTHVINDLGLQSSCTMVV 508  
QY 159 PDVSTLPHFNQKDE-----PKMRLVAMLYTPPLRSNGSGDDVTVSCRIILTRSP 208  
Db 509 FWISNVTVRQTQDSFTSGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN----- 558  
QY 209 EFDFTYLVPPPTVESKTKPFTPLVLTGELSISRPPSLSDMVTSPNESIVVQPNQGRVTL 268  
Db 559 DFSVRLLEDIT-----HISQALPQGIEDLT-----SEVAQ----- 589  
QY 269 DCELLGTTLOACNICSIRGKVTGOVPSEQHMMNLEITNLNGTQDFDTPDDVPA----- 321  
Db 590 -----GALTSLPKQD--SLPDTKASGPAH--SKEVPALTAVETG 626  
QY 322 ---PLGVDPDFAGEVGVLSQRNGES-----NPNRAHDVAV--- 355  
Db 627 ATNPL-APSDTVQTRHVHVQRSSRSESTIESPFARGACVATIEVDNEQPTTQAQKLFAWVR 685

QY 356 ATYSKYTPKGLVIGTWTNDVENQTKETPI---GLNEVANGHREPOWTLPRYS GAL 412  
DB 686 ITYKDTVQLRRKL-EFTYSRFDME-----FTFVVTANFTNANGHALNQ----- 729  
QY 413 TLNMLAPAVAPLFGERLLFFRSYVPLKGGFNPALDCSPQEWVQHFYQESA-PSL 469  
DB 730 -----VYQIMYIPP-----GAP-----TPKSWDDYTWTQTSNPSI 759

RESULT 11  
US-08-461-503-2  
; Sequence 2, Application US/08461503  
; Patent No. 5834302  
; GENERAL INFORMATION:  
; APPLICANT: Racaniello, Vincent  
; APPLICANT: Tatem, Joanne M.  
; APPLICANT: Weeks-Levy, Carolyn L.  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES  
; TITLE OF INVENTION: FROM CDNA  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10112  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/461,503  
; FILING DATE: 5-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; TELEFAX: 422523 COOP UI  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2206 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-461-503-2

Query Match 3.9%; Score 111.5; DB 2; Length 2206;  
Best Local Similarity 17.9%; Pred. No. 0.47;  
Matches 107; Mismatches 170; Indels 247; Gaps 28;

QY 11 SNDGAAGLVP-----ESNNEAMALEPV-----VGASLAAPVT----- 42  
DB 270 TNSATIVLPYVNALAIDSMVKHNNWGIAILPLSLDFAQDSSVEIPITVTIAPMCSEFN 329  
QY 43 GQTNIDP-----WIRTFVQAPNG--EFTVSPR-NSPGEI----- 75  
DB 330 GLRNVTPAKFQGLPVLNTPGNSQYLTSDNHQSPCAIPEDFTVPIDIPGEVKNMMELAEI 389  
QY 76 -----LVNLELGPENLPYLAH-----LARMYNG 98  
DB 390 DTMIPLNLESTKNTMDMYRVLTLSADLSQPICLSLSPAFDPRLSHTMLGVLNLYTH 449  
QY 99 YAGGMEVQVWLAGNAFTAGKIIFAAVPPYFVENLSPSQITWFPFHVIIDVRTLPEVLLPM 158  
DB 450 WAGSLKFTFLFCGSMMATGKILVAYAPGAQPPT-SRKEAMLGTHVINDLGLQSSCTMWV 508

QY 159 PDVRSITLPHFNQDE-----PKMRLVAMLYTPLRSNGSGDVFVTVSCRILTRPSP 208  
DB 509 PMISNVYRQTQDSFTGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN----- 558  
QY 209 EFDFTYLVPEPTVESKTFPLVLTLGLBSNRRPPLSIDEMWTSNPSIIVQVQNGRVTL 268  
DB 559 DFSVRLRLRTT-----HISQALPQGIEDLT-----SEVAQ----- 589  
QY 269 DGLLGTTLQACNICSRGKVTQVPSQEQHMWNLEITNLNGTQDPDTPDDVPA----- 321  
DB 590 -----GALTSLPKQOD---SLPDKASGPAH--SKEVPALTAVETG 626  
QY 322 ---PLGVDPFAGEVFCVLSQRNRGES-----NPNARHADAVV--- 355  
DB 627 ATNPL-APSDTVQTRHVQVRRSRSESTIESFFARGACVAIEVDNEQPTTTRAKLFAAMWR 685  
QY 356 ATYSKYTPKGLVIGTWTNDVENQTKETPI---GLNEVANGHREPOWTLPRYS GAL 412  
DB 686 ITYKDTVQLRRKL-EFTYSRFDME-----FTFVVTANFTNANGHALNQ----- 729  
QY 413 TLNMLAPAVAPLFGERLLFFRSYVPLKGGFNPALDCSPQEWVQHFYQESA-PSL 469  
DB 730 -----VYQIMYIPP-----GAP-----TPKSWDDYTWTQTSNPSI 759

RESULT 12  
US-08-465-250-2  
; Sequence 2, Application US/08465250  
; Patent No. 6136570  
; GENERAL INFORMATION:  
; APPLICANT: Racaniello, Vincent  
; APPLICANT: Tatem, Joanne M.  
; APPLICANT: Weeks-Levy, Carolyn L.  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM  
; TITLE OF INVENTION: CDNA  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release 1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,250  
; FILING DATE: 6-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; TELEFAX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2206 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-465-250-2

Query Match 3.9%; Score 111.5; DB 3; Length 2206;  
Best Local Similarity 17.9%; Pred. No. 0.47;  
Matches 107; Mismatches 170; Indels 247; Gaps 28;

QY 11 SNDGAAGLVP-----ESNNEAMALEPV-----VGASLAAPVT----- 42

Db 270 TNSATIPYVNALAIDSMVRHNNWGIAILPLSPLDFAQDSVEIPITVTIAPMCSEFN 329  
QY 43 GQTNIIIDP-----WIRTFVQAPNG--EFTVSPR-NSPGEI-----75  
Db 330 GLRNVTAQFQGLPVLTGSGNQYLTSDNHQSPCAIPEFDVTPPIDIPGEVKNMELAEI 389  
QY 76 -----LVNLELQPELNPYLAH-----LARMYNG 98  
Db 390 DTMIPLNLESTKRNTMDMYRVLTSDSADISQPIFCLSLSPAFDPRLSHTMLGEVLNYTH 449  
QY 99 YAGGMEVQVNLGNAFTAGKILFAAVPPVFPVENLSPSQITWFFPHVIIDVRTLEPVLPM 158  
Db 450 WAGSLKFTFLFCGSMWATGKILVAYAPGAQPPPT-SRKEAMLGTHVIMDLGQSSCTMVV 508  
QY 159 PDVRSTLFHFNQKDE-----PKMLVAMLYTPLRNSGDDVFTVSCRILTRPSP 208  
Db 509 PWISNVYRQTQDSFTEGGYISMFYQTRIVVPLSTPKSMMLG---FVSACN-----558  
QY 209 EFDFTVLVPTVESKTKPFTPLVLTGELSNRFPPLSIDEMVTSNPNESIVVQPPQNGRVTL 268  
Db 559 DFSVRLLRDTT-----HISQALPOGIEDLT-----SEVAQ-----589  
QY 269 DGEELGTTLOACNCSIRGKVTGQVPSQEHMWNLEITNLNGTQDPTDDVA-----321  
Db 590 -----GALTSLPKQOD--SLPDTTKASGAH--SKEVPALTAVETG 626  
QY 322 ---PLGVPDPAGEVGLVLSQRNGES-----NPNRAHDAVV---355  
Db 627 ATNPL-APSDTVQTRHVVRORSESTIESFFARGACVAIIEDNEQPTTRAKLFAMWR 685  
QY 356 ATYSDKYTPKGLVQIGTWTNDVENVQTKFTPI---GLNEVANGHRFQWTLPRYS GAL 412  
Db 686 ITYKDTVOLRRKL-EPFTYSRDEME-----FTFVVTANFTNANGHALNQ-----729  
QY 413 TLNMLAPAVAPLFPGERLLFRSYVPLKGGNPAIDCSVPQEWQHFYQBSA-PSL 469  
Db 730 -----VYQIMYIPPP---GAP-----TPKSWDDYTWTQSSNPISI 759

## RESULT 13

US-09-513-783A-152

; Sequence 152, Application US/09513783A

; Patent No. 6416959

; GENERAL INFORMATION:

; APPLICANT: Giuliano, Kenneth A.

; APPLICANT: Kapur, Ravi

; TITLE OF INVENTION: A System for Cell Based Screening

; FILE REFERENCE: 97-022-L1

; CURRENT APPLICATION NUMBER: US/09/513,783A

; CURRENT FILING DATE: 2000-02-25

; NUMBER OF SEQ ID NOS: 180

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 152

; TYPE: PRT

; ORGANISM: Mus musculus

US-09-513-783A-152

## Query Match

Best Local Similarity 3.8%; Score 108.5; DB 4; Length 1125;

Matches 130; Conservative 65; Mismatches 232; Indels 187; Gaps 29;

QY 5 SNDATPSNDGAGLVPESSNNEAMALEPVVVGASLAAP-----VTG-----QT 45  
Db 256 STDTTPGPDTEAALAKDI--EITKPDVILANVTQSTESDMFLQADMELLTGTEAAHAN 313  
QY 46 NIIDPW-----IRTNFVQAP-----NGEFTVSPRNSPGIILVNLELGP-----EL 85  
Db 314 NIILPTEPDESSTKDVAPPMBEIEIVPGNDTTSFKETETTLPIKMDLAPPEDVLLTKETEL 373  
QY 86 NPYLAHLARMYNGYAGCM-----EVQVMLAGNAFTAGKIIFAAVPPVFPVENLSPSQITW 141

Db 374 AP-----AKGMVLSLSEIEALAKNDVRSABE1-----PVAQETVVSETEVLA 415  
QY 142 PHVII---DVRTL-EPVLLPMPDVRSTLTFHFNQKDEPKMLVAMLYTPLRSH- GSGDDVF 196  
Db 416 TEVVLPSDPIITTLTKDVTLPLEAERPLVDTMTPLSLETETMTLCKETAPPTETNLGMAKDM- 474  
QY 197 TVSCILIRPSPPEPFT-----YLVPPPTVESKTKPFTPLVLTGELSNRFPPLSIDEMVT 251  
Db 475 -----SPLPESEVTILGKDVVILP---ETKVABF-----NNVTPLS-EEVET 511  
QY 252 SPNESIVVQPPQNGRVTLGELIGTLTQLOACN-----ICSIKGVKTVGOVPS 296  
Db 512 SYKDNPSAEYEPAPLAKNADLHSGTELVINDSNMAPASDLALPLETKVATVPKDKGTVOT 571  
QY 297 EOHMWNLEITNLNGTQDPTDDVPAPLGVDP--PAGEVFGV-----LSQRNRGE 343  
Db 572 BEK--PREDSQLASHQHKQSTVPPCTASPEPVKAAEQMSTLPIDAPSPLENLEQKETPG 629  
QY 344 SNPA-----NRAHDAVATYSDKYTP-----KGLVQIGTWTNDVE 380  
Db 630 SQSPSPSGVSRQEEBAKAAVGTGNDITTPPNKPEPPSPKAKPLATTQPAKTSTSKAK 689  
QY 381 NQPTKFTPIGLNEVANGHRFQWTLPRYSGALTLN-MNLAPAVAPLFPGER-----LLPF 434  
Db 690 TQPTS-----LPKQAPATTSGGLNKKPMSLASGVSVAAPHKRPAAATATAR 735  
QY 435 RSYVPLKGGFGNPAIDCSVPQEWQHFYQBSAPSLGDVALVRYVNPDTGRVLFPEAKLHKG 494  
Db 736 PSTLPARDVKPKPITEAKVAEKRTSPSPSSAPALKP-----GPKTTPTVSKA-----783  
QY 495 GELTVSST--STGP 506  
Db 784 ---TSPSTLVSTGP 794

## RESULT 14

US-09-513-783A-22

; Sequence 22, Application US/09513783A

; Patent No. 6416959

; GENERAL INFORMATION:

; APPLICANT: Giuliano, Kenneth A.

; APPLICANT: Kapur, Ravi

; TITLE OF INVENTION: A System for Cell Based Screening

; FILE REFERENCE: 97-022-L1

; CURRENT APPLICATION NUMBER: US/09/513,783A

; CURRENT FILING DATE: 2000-02-25

; NUMBER OF SEQ ID NOS: 180

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 22

; TYPE: PRT

; LENGTH: 1610

; ORGANISM: Artificial Sequence

; OTHER INFORMATION: Description of Artificial Sequence:

; OTHER INFORMATION: EYFPP-DEVD-MAP4-EBFP construct

US-09-513-783A-22

## Query Match

Best Local Similarity 3.8%; Score 108.5; DB 4; Length 1610;

Matches 130; Conservative 65; Mismatches 232; Indels 187; Gaps 29;

QY 5 SNDATPSNDGAGLVPESSNNEAMALEPVVVGASLAAP-----VTG-----QT 45  
Db 502 STDTTPGPDTEAALAKDI--EITKPDVILANVTQSTESDMFLQADMELLTGTEAAHAN 559  
QY 46 NIIDPW-----IRTNFVQAP-----NGEFTVSPRNSPGIILVNLELGP-----EL 85  
Db 560 NIILPTEPDESSTKDVAPPMBEIEIVPGNDTTSFKETETTLPIKMDLAPPEDVLLTKETEL 619  
QY 86 NPYLAHLARMYNGYAGCM-----EVQVMLAGNAFTAGKIIFAAVPPVFPVENLSPSQITW 141  
Db 620 AP-----AKGMVLSLSEIEALAKNDVRSABE1-----PVAQETVVSETEVLA 661

us-09-926-799-7.rai

Wed Jun 2 09:13:32 2004

```

QY 142 PHVII--DVRTL-EPVLLPMPDVRSTLHFHFNQXDEPKRILVAMLYTPLRSN-GSGDDVF 196
Db 662 TEVVLPSDPTITLTLDVTLPLEAERPLVDTMTPTSLTEMTLGKTAFTPTETNLGAKDM- 720
QY 197 TVSCRILTRPSPEFDT-----YLVPPTVESKTKPFTLPLVLTGELNSRPLSIDEMVT 251
Db 721 -----SPLSEVTLGKDVILP---ETKVAEF-----NNVTPLS-EEBVT 757
QY 252 SPNESIWPQNGRVTLDGELLGTTQLOACN-----ICSIRGKVTGOVPS 296
Db 758 SVKDMSPSAETEAPLAKNADLSGTSLIYDNSMAPASDLAULETKVAIVPIKKGTVQT 817
QY 297 EGHMWNLEITNLTNGTOFTDDVPAPLGVPD--FAGEVFGV-----LQQRNKE 343
Db 818 EEK--PREDSQLASMHQKQSTVPCTASPEPKAEQMSLIPIDAPSPLENLEKETPG 875
QY 344 SNPA-----NRAHDADVATYSKYTP-----KGLVQIGTWNNDVE 380
Db 876 SQFSEPCGVSQRQEEAKAAVGTGNDITTPPNKEPPSPKAKPLATTQPAKTSTSKAK 935
QY 381 NQTKFTPIGLNEVANGHREFTQWTLPRYSALTIN-MNLAPAVAPLFFGER-----LLFF 434
Db 936 TQPTS-----LPKQAPTTSGLNKKPKMSLASGVPAPAPKRPAAATATAR 981
QY 435 RSVVPLKGGFGFNPAIDCSVPQVQHFYQESAPSLGDBVALRVNPDTCGRVLFEAKLHG 494
Db 982 PSTLPARDVKPKPITEAKVAEKRTSPSKPSAPALKP-----GPKTTPTVSKA----- 1029
QY 495 GFLTVSST--STGP 506
Db 1030 ---TSPSTLVSTGP 1040

```

RESULT 15

```

US-09-489-039A-13081
; Sequence 13081, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; PRIOR FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 13081
; LENGTH: 776
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-13081

```

Query Match 3.7%; Score 105.5; DB 4; Length 776;  
 Best Local Similarity 21.2%; Pred. No. 0.34;  
 Matches 120; Conservative 63; Mismatches 218; Indels 165; Gaps 26;

```

QY 19 VPESNNEAMALEPV-----VGASLAAPVTGQNIIDPWIRTNFVQAPNGEFTVSPRN 70
Db 24 VESSEGEIVAVKVPWPKNPSRIGSLQSDPAVTSQTRIRPAPVAGYLQ--HGPVSGREG 81
QY 71 SPOEILVNLELGPENLPYLALHAR-----MYNGYAG-----GMEVOV 107
Db 82 KEPFVEVSEVALDL-----LARELSVKARCONAEIYGSYGWASAGRPHHAQSLHR 135
QY 108 MLAG-NAFTAGKIIFAAVPPYFPVENLSPSQMTFPHVITDVRTLEPULLPMPDYRSTLF 166
Db 136 FLKGGGYTASTNTYS-----SAAGERILPH-----ILGPLSLPHROHT 174
QY 167 HFNQKDEPKRILVAMLYTPLRS--NGSG-----DDVFTVSCRILTRPSPEFDT 213
Db 175 HFSELARECOLFVAIGGLPLRNAQVNGGGANDHMLQWLDKQOANGTREFI-NISFVRNDL 233

```

```

QY 214 YLVPTVESKTKPFTLPLVLTGELNSRPLSIDEMVTSPNESIWPQNGRVTLDGELL 273
Db 234 SAVPDAEWLAIRPGTDTALLLA-LSYVLIABSLYDQAFVASHTVGFAPY--RAYLLGEHD 290
QY 274 GT-----TQLOACNIC-----SIRGKVTGOVPSBOHWW-NLEITNLN 309
Db 291 GVAKTPAMAAAITGLDAQRIADLAREMARHRTMWNISIQARQGEQAYWATVALL 350
QY 310 GTQFDPTDDVPAPLGVDPFAGEVFGVLSQNRNGESNPA-NRAHDVAV--VATYSKYTKPKL 366
Db 351 GQLGTGGGLGFGYACTNLGAVRKAFS-----GRLPAGENAVDSVIPVARLSDM----- 401
QY 367 GLVQIGTWNNDVENOPTKFTPIGLNEVANGHRF-----EQWTLPR-----Y 408
Db 402 -LLHPG--ETYPFDGQORYPDIRLVYWAGGNAFHHQDINRLCEAWRRPPTVVVHEQYM 458
QY 409 SGALTILNMLAPAVAPLFFGER-----LLFFRSYVPLKG----- 442
Db 459 TAQAKFSDIPLPATTSL---EREDIGSGHGDGFMIAMSAQIPVGEARDDYAFCDLAGR 515
QY 443 -GFGNPAIDCSVPQVQHFYQESAP 467
Db 516 LGFGEAFSEGRDAGQWLRLHYESRP 541

```

Search completed: June 1, 2004, 13:58:02  
 Job time : 14.8658 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.5201 Seconds  
(without alignments)  
4368.312 Million cell updates/sec

Title: US-09-926-799-7  
Perfect score: 2845  
Sequence: 1 MKMASNDTPSNDGAGLVP.....VNQFYSLAPMTGTGNGRRRVQ 540

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1233	43.3	530	14	US-10-314-739-3
2	296.5	10.4	668	14	US-10-209-507-4
3	292.5	10.3	547	12	US-10-670-695-4
4	292.5	10.3	671	12	US-10-670-695-2
5	273	9.6	669	14	US-10-209-507-2
6	138	4.9	1147	15	US-10-327-481A-38
7	131.5	4.6	338	10	US-09-867-932A-2
8	117.5	4.1	5376	15	US-10-028-248A-74
9	117.5	4.1	5376	15	US-10-107-782-74
10	111	3.9	6304	14	US-10-147-026-16
11	109.5	3.8	689	14	US-10-207-706-2
12	109	3.8	257	12	US-10-282-122A-76092
13	108.5	3.8	1125	14	US-10-100-957A-152
14	108.5	3.8	1521	14	US-10-238-075-909
15	108.5	3.8	1539	14	US-10-085-959-253

16	108.5	3.8	1610	14	US-10-100-957A-22
17	108.5	3.8	5374	15	US-10-028-248A-75
18	108.5	3.8	5374	15	US-10-107-782-75
19	107.5	3.8	1152	15	US-10-369-493-12027
20	107	3.8	1544	15	US-10-369-493-6078
21	106.5	3.7	882	14	US-10-097-340-37
22	106.5	3.7	894	12	US-10-435-114-39106
23	106.5	3.7	561	12	US-10-282-122A-52830
24	104.5	3.7	627	9	US-09-862-027-46
25	104.5	3.7	968	14	US-10-101-464A-76
26	103.5	3.6	785	15	US-10-369-493-11883
27	103.5	3.6	3930	12	US-10-282-122A-46817
28	102.5	3.6	1501	12	US-10-282-122A-44608
29	101.5	3.6	878	9	US-09-905-983-48
30	101.5	3.6	878	13	US-10-165-049-2
31	101.5	3.6	878	13	US-10-165-049-3
32	101.5	3.6	878	15	US-10-173-551-34
33	101.5	3.6	2468	12	US-10-282-122A-66335
34	101.5	3.6	2468	14	US-10-246-330-4
35	101	3.6	609	15	US-10-369-493-22725
36	101	3.6	609	15	US-10-369-493-22806
37	101	3.6	925	15	US-10-108-260A-3889
38	101	3.6	6310	12	US-10-282-122A-67793
39	100.5	3.5	903	15	US-10-369-493-14300
40	100.5	3.5	1147	15	US-10-369-493-11450
41	100.5	3.5	1151	15	US-10-369-493-14817
42	100.5	3.5	1151	15	US-10-369-493-15006
43	100.5	3.5	2358	12	US-10-282-122A-45763
44	100	3.5	578	9	US-09-924-358-39
45	100	3.5	578	15	US-10-410-764-39

ALIGNMENTS

RESULT 1  
US-10-314-739-3  
Sequence 3, Application US/10314739  
Publication No. US20030129588A1  
GENERAL INFORMATION:  
APPLICANT: Estes, Mary K  
Jiang, Xi  
Graham, David Y  
TITLE OF INVENTION: Methods and Reagents to Detect and Characterize No. US20030129588A1walk and Related Viruses  
NUMBER OF SEQUENCES: 4  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Pubbright & Jaworski L.L.P.  
STREET: 801 Pennsylvania Ave., N.W.  
CITY: Washington, D.C.  
STATE: <Unknown>  
COUNTRY: USA  
ZIP: 20004  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/314,739  
FILING DATE: 09-Dec-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/486,049  
FILING DATE: June 7, 1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Davis, Peter  
REGISTRATION NUMBER: 36,119  
REFERENCE/DOCKET NUMBER: 311.023  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-662-0200  
TELEFAX: 202-662-4643  
TELEX: <Unknown>

```

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-209-507-4

Query Match          10.4%; Score 296.5; DB 14; Length 668;
Best Local Similarity 26.0%; Pred. No. 8.1e-19;
Matches 114; Conservative 55; Mismatches 135; Indels 135; Gaps 22;

QY 11 SNDGAAGLVPESSNEAMALEPVGASLAAPVTGQTNIIIDPWIRTNFVQAPNGEFTV--- 66
DB 125 ADDGSV-TTPEOQT-----LVGGVIAEPNAQMSAVAD--VATG--KSVSEWEAFPSF 172
QY 67 -----SPRNSGPEILVNLELGPENLPYLALHARMYNGVAGMEVQVMVQVLAGNAFTAGKII 120
DB 173 HTSVNMTSETQKILFKQSLGPLNLPYLTHLAKLYVAWSGSIEVRFSISGSGVFGKLA 232
QY 121 FAAVPPVF-PVENLSPSQITMFPFHVIIDVRLTLEPVLPMVDVRSITLPHFNQKDEPKMRLV 179
DB 233 AIVVPGCIDPVQSTSMQ--YPHVLFDAQVPEPVIPTIDLRNLSLYHL-MSDITTTSLV 288
QY 180 AMLYTPI-----RNSGSGDDVFTVSCRILTRPSPEFDFTYLVPPTVESKTKPFTLPVLJT 233
DB 289 IMIYNDLINFYANDSNSSG-CIVTVE---TKGPDFKFKHLKPPG-----SMLT 333
QY 234 LGELNSRRFPLSIDEMVTSNESIVVQPNQGRVTLTGELLGTTLOACNICSRGKVTGQ 293
DB 334 HGS1-----PSDLI----- 342
QY 294 VPSQHMWNLEITNLNGTQDPTDDVPAPLGVDPDFAGEVGV-----LSQRNRGESNPAN 348
DB 343 -PKSSSLM--IGNRHS--DITDFVIKPP-----VFQANRHFDFNQETAGWSTPRF 388
QY 349 RAHDAAVATYSKYTPKGLVQIGT-----MNTNDVENQPTKFTPIGLNEVANGH 398
DB 389 R---PITTVSEKSGSKLG-IGVATDSIVPGIPDGWPDITI---PEKLTAGDYAITNGG 441
QY 399 RFEQWTLPRYSGALTNNN 417
DB 442 NNDITTAADYDGAIIKKN 460

RESULT 3
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication No. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; TITLE OF INVENTION: STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 547
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-670-695-4

Query Match          10.3%; Score 292.5; DB 12; Length 547;
Best Local Similarity 22.7%; Pred. No. 1.4e-18;
Matches 133; Conservative 84; Mismatches 223; Indels 147; Gaps 25;

QY 11 SNDGAAGLVPESSNN--EAMALEPVGASLAAPVTGQTNIIIDPW-IRTNFVQAPNGEFTVS 67
; NUMBER OF SEQ ID NOS: 26

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match          43.3%; Score 1233; DB 14; Length 530;
Best Local Similarity 47.6%; Pred. No. 3e-108;
Matches 263; Conservative 81; Mismatches 163; Indels 46; Gaps 11;

QY 1 MKAASNDATPSNDGAAG---LYPESN-NEAMALEPVGASLAAPVTGQTNIIIDPWIRTNF 56
DB 1 MMMAASNDATSSVDGASGAOLPVEVNASDPLAMDVPAGSTAVATAGVNPIDPWIIINF 60
QY 57 VQAPNGEFTVSPRNSGPEILVNLELGPENLPYLALHARMYNGVAGMEVQVMVQVLAGNAFTA 116
DB 61 VQAPQGEFTISPNNTGDLVFLDLSGLPHLNPFLHLSQMYNGWGMVRIMLAGNAFTA 120
QY 117 GKIIFAAVPPYFVVENLSPSQITMFPFHVIIDVRLTLEPVLPMVDVRSITLPHFNQKDEPKM 176
DB 121 GKIIICIPPGFSGHNLITAAQATLFPHVIADVRITLDFIEVPLEDVNVLPHNDRNQTM 180
QY 177 RLVAMLYTPLRNSGSGDDVFTVSCRILTRPSPEFDFTYLVPPTVESKTKPFTLPVLTLGE 236
DB 181 RLVCMLYTPLRNGTGTGDSFVAGRVMTCPSPDFNLFVLPPTVEQKTRPFTLPNPLSS 240
QY 237 LNSRRFPLSIDEMVTSNESIVVQPNQGRVTLTGELLGTTLOACNICSRGKVTQVPS 296
DB 241 LNSRAPLPISSMGISSPNQSVQFQNGRCRLDGRVLGTTIPVSLSHVAKIRGTSNCTV-- 298
QY 297 EQHMWNLEITNLNGTQDPTDDVPAPLGVDPDFAG-----EVFGLVSQRNRGESNPANR 349
DB 299 -----INLTEDGTTPHPFEG-PAPIGFPLDGLSCDWHINMTQFGHSSQTQ----- 342
QY 350 AHDAAVATYSKYTPKGLVQ---IGTWNNDVENQPTKFTPIGLNEVANGHREFEOWTLTP 406
DB 343 -YD--VDTPDTFVPHLGSIQANGTSGNYGV-----LSWISPPSHPSGSQVDLMKIP 393
QY 407 RYSGALTLMNLAPAVAPLFGERLLPFRSVVPLKGGFGNPAIDCSVPQEWQHFYQESA 466
DB 394 NYGSSITRATHLAPSVYPPGGEVLVFFWKMPPGEGAINLP---CLLPQEVYISHLASEQA 450
QY 467 PSLGVDVALVRYNPDTRGLVFEAKLHKGGLFTV--SSTSTGTVVVPANGYFKEDSWNQF 524
DB 451 PTVGEAALLHYVDPTGRNLGFKAYPDGFLTCVPNGASSGQQPLPINGVFVFSWVSRRF 510
QY 525 YSLAPMGTGNGRR 537
DB 511 YQLKPVGTASSAR 523

RESULT 2
US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
```



```
Db      1  ADDGSI-TAPEOGTVMGVIAEPPSQMSTAADMATGKSDVSEWEAFFSHTSVNM-----S 55
Qy      68  PRNSPGEILVNLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAGTAGKIIFAAVPPY 127
Db      56  TSETQKILFKQSLGPLNPLYHLAKLYAVWAGSIEVRFSISGSGVGGKLAIVVPP- 114
Qy     128  FPVENLSPSQITMFPVHIIDVRLTPVLLPMPDVRSTLPHFNQKDEPKMRLVAMLYTPLR 187
Db     115  -GVDVPQSTMLQYPHVLFDAQVPEVIFCLPDLRLSTLYHL-MSDTRTSLVIMVYNDL- 171
Qy     188  SNGSGDDVFTVSC--RILTRSPPEPDEFTYLVPP-----TVESKTKPPTLPLVLTGELS 238
Db     172  INPYANDANSSGCIYTVETKPGDPFKHLLKPPGSMMLTHGSIPLSOLIPTK-SSLWIG--- 227
Qy     239  NSRFLPSIDEMVTSP-----NE-----SIVVQPQNGR----- 265
Db     228  -NRYWSDITDFVIRPFVQANRHFDFNQETAGMSTPRPRPISVITTEQNGAKLGIGVATD 286
Qy     266  -----VTLDGELLGTTQLQACN-----ICSIKRGKVTQV-----P 295
Db     287  YIVGIPDGMPTTIPGELIPAGDYAITNGTGNDDITTATGYDTADIKNNTNFRGMWICG 346
Qy     296  SEQHWNLEITNLNGTQDPTDDVPAPLGVDPDFAGEVGVLSQRNGESNPANRAHDVAV 355
Db     347  SLQRAWGDK--KISNTAFITT-----ATLDGDNNNKINPCNTIDQSKI 387
Qy     356  ATYSKYTPKLGVLQIGTWTNTNDVENQPTKFTPIGLNEVANGHREFQ-----WTLPRYSCA 411
Db     388  VVFQDNHVGKKA-----QTSDDTLALLGYT--GIGEAIGSDRDRVVRISTLPE-TGA 437
Qy     412  LTLNNLAPAVAPLPPGERLLFFRSYVPLKGGFNPADICSPQEWVQHFYQESAPSLGD 471
Db     438  R-----GGNHPIFYKNSIKL--GVYIRSIDV-----FNSQILHTSRQ 472
Qy     472  VALVRYVNPDTGRVLFPEAKLHKGGLFVSSTSTGPPVVPANGYFKFED 518
Db     473  LSLNHYLLPPDSFAVYRIIDNSNGWFIDIGSDGFSFVGSGFGKLE 519
```

## RESULT 4

```
US-10-670-695-2
; Sequence 2, Application US/10670695
; Publication NO. US20040058316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; TITLE OF INVENTION: STATUS OF AN ANIMAL
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; PRIOR FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 2
; LENGTH: 671
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-670-695-2
```

```
Query Match      10.3%; Score 292.5; DB 12; Length 671;
Best Local Similarity 22.7%; Pred. No. 28-18;
Matches 133; Conservative 84; Mismatches 223; Indels 147; Gaps 25;

Qy     11  SNDGAGLVPESSN--EAMALPVPVGCASLAAPVTGOTNIIDPW-IRTNFVQAPNGEFTVS 67
Db     125  ADDGSI-TAPEOGTVMGVIAEPPSQMSTAADMATGKSDVSEWEAFFSHTSVNM-----S 179
Qy     68  PRNSPGEILVNLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAGTAGKIIFAAVPPY 127
```

```
Db     180  TSETQKILFKQSLGPLNPLYHLAKLYAVWAGSIEVRFSISGSGVGGKLAIVVPP- 238
Qy     128  FPVENLSPSQITMFPVHIIDVRLTPVLLPMPDVRSTLPHFNQKDEPKMRLVAMLYTPLR 187
Db     239  -GVDVPQSTMLQYPHVLFDAQVPEVIFCLPDLRLSTLYHL-MSDTRTSLVIMVYNDL- 295
Qy     188  SNGSGDDVFTVSC--RILTRSPPEPDEFTYLVPP-----TVESKTKPPTLPLVLTGELS 238
Db     236  INPYANDANSSGCIYTVETKPGDPFKHLLKPPGSMMLTHGSIPLSOLIPTK-SSLWIG--- 351
Qy     239  NSRFLPSIDEMVTSP-----NE-----SIVVQPQNGR----- 265
Db     352  -NRYWSDITDFVIRPFVQANRHFDFNQETAGMSTPRPRPISVITTEQNGAKLGIGVATD 410
Qy     266  -----VTLDGELLGTTQLQACN-----ICSIKRGKVTQV-----P 295
Db     411  YIVGIPDGMPTTIPGELIPAGDYAITNGTGNDDITTATGYDTADIKNNTNFRGMWICG 470
Qy     296  SEQHWNLEITNLNGTQDPTDDVPAPLGVDPDFAGEVGVLSQRNGESNPANRAHDVAV 355
Db     471  SLQRAWGDK--KISNTAFITT-----ATLDGDNNNKINPCNTIDQSKI 511
Qy     356  ATYSKYTPKLGVLQIGTWTNTNDVENQPTKFTPIGLNEVANGHREFQ-----WTLPRYSCA 411
Db     512  VVFQDNHVGKKA-----QTSDDTLALLGYT--GIGEAIGSDRDRVVRISTLPE-TGA 561
Qy     412  LTLNNLAPAVAPLPPGERLLFFRSYVPLKGGFNPADICSPQEWVQHFYQESAPSLGD 471
Db     562  R-----GGNHPIFYKNSIKL--GVYIRSIDV-----FNSQILHTSRQ 596
Qy     472  VALVRYVNPDTGRVLFPEAKLHKGGLFVSSTSTGPPVVPANGYFKFED 518
Db     597  LSLNHYLLPPDSFAVYRIIDNSNGWFIDIGSDGFSFVGSGFGKLE 643
```

## RESULT 5

```
US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-2
```

```
Query Match      9.6%; Score 273; DB 14; Length 669;
Best Local Similarity 22.4%; Pred. No. 1.4e-16;
Matches 120; Conservative 74; Mismatches 188; Indels 154; Gaps 23;

Qy     67  SPRNSPGHILVNLELGPENLPYLAHLARMYNGYAGGMEVQVMLAGNAGTAGKIIFAAVPP 126
Db     180  STSETQKILFKQSLGPLNPLYHLAKLYAVWAGSVDVRFSGSGVGGKLAIVVPP 239
Qy     127  YFPVENLSPSQITMFPVHIIDVRLTPVLLPMPDVRSTLPHFNQKDEPKMRLVAMLYTPL 186
Db     240  --GVDVPQSTMLQYPHVLFDAQVPEVIFSIPLRLSTLYHL-MSDTRTSLVIMVYNDL 296
```

4



```

Db 797 PEE-TSVPTETISTEVTVSPETTLTPTVPTVSTVTVNVSPTSVPPEETILTLYT 855
QY 185 -----PLRNSGDDVFTVSCRILTRPSPEFDTYLVVPVTSKTKFTPLVLTGLBSLN 239
Db 856 EYPTVPTVTVGHTVTVNVSPTSVPTEETISTEVTVSPETTLTPTVPTVST-EVTN 914
QY 240 SRPPLSIDEMVTSNPSIV-----VQPQNGRVTLDGELLGTTQLQACNICSIRGKVTGQV 294
Db 915 -----VSPETSVPPEETILTTEITVSPBETVPPIEGTLPTFEVL-----TV 956
QY 295 PSEQHMMNLEITNLNGTQDPTDDVPAPLGVDFAGEVGV 335
Db 957 P-----IEVTTF-----PTGETVPTVPTVSTEMTGV 984

RESULT 10
US-10-147-026-16
; Sequence 16, Application US/10147026
; Publication No. US20030003538A1
; GENERAL INFORMATION:
; APPLICANT: Dietrich, Paul Shartzer
; APPLICANT: Huang, Chiao-Chain
; APPLICANT: Johnson, Carl D.
; APPLICANT: Sangameswaran, Lakshim
; TITLE OF INVENTION: Neuropathic Pain Genes; Compositions
; TITLE OF INVENTION: Thereof; Related Reagents
; FILE REFERENCE: ROCH-006
; CURRENT APPLICATION NUMBER: US/10/147,026
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/155,702
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/189,931
; PRIOR FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 6304
; TYPE: PRT
; ORGANISM: human
US-10-147-026-16

Query Match
Best Local Similarity 3.9%; Score 111; DB 14; Length 6304;
Matches 100; Conservative 77; Mismatches 168; Indels 172; Gaps 25;

QY 42 TQGTNIIDPWIRTNVQA-----PNEGFTVSPRSPGEILVNLELGPENLPHLAARMY 96
Db 3884 TQPSVRRPGMEIAEIMTEENDDPRIEMFHVTRGAGEVITAYEVPPPLNV----- 3934
QY 97 NGYAGMEVQVWMLAGNFTAGKIIFAAVPPYPPFVENLSPSQTITWPFHVIDVRLTLEPVL 156
Db 3935 -----LQPVVRLAGSFGANVYWKAGSDSAGLEDFFPS-----HGLEFADKQVTAM 3982
QY 157 PMPDVRSTLTFHFNQDEPKRVLVAMLYTPLRSNGS-----GDDVFTVSCRILTRPSPEFDT 212
Db 3983 -----IITII-----DDAEFELTETNTSLISVAGGRLGDDV-----VTVIPIQND- 4026
QY 213 TVLVPTVTSKTKFTPLVLTGLBSNRPPLSIDEMVTS-----PNESI-----VVQPQNGRVT 267
Db 4027 -----SPF-----GVFGFEKTMIDESLSDDDPSVTILTVVRSFGK----- 4065
QY 268 LDGELLGTTQLQACNICSIRGKVTGQVSPSEHMMNLE-----ITNLNGT-QFDPTDDVP 320
Db 4066 -----GTVRL-----WTIDEKAKNLSPLNGTLHFDETE----- 4095
QY 321 APLGVDFAGEVGVLSQNRNGESNPANRAHDVAVATYDKYTPKGLVQIGTWTNDVE 380
Db 4096 -----SQKT-----IVLHTQVTLVEEDRRFTIQLISIDEVE 4127
QY 381 NQPTKFTP-----IGLNEVANG-----HRFEQWTLPRYSGLTLMNMLAPAVAPLFP 428
Db 4128 ISPVKGSASIIIRGDKKASGEVGIAPSSRILIGESAKYNGTAIISLVRGAIL-----G 4183

```

```

QY 429 ERLLEFRSYVLKGGFGNPAIDCSVPOEWHQHFQESAPSLGDLVALRVYVNPDP--TGRVL 486
Db 4184 EYTVFWRIFPPSVGSEFASTSKLTMRE-----QNAV-----IIVIQALNDDIPEEKSF 4232
QY 487 PEAKL-----HKGGLFVTSSTSTGPPVVVPAN-----GYFKF 517
Db 4233 YEFQLTAVSEGGVLSSESSSTANITVWASDSPYGRFAF 4269

RESULT 11
US-10-207-706-2
; Sequence 2, Application US/10207706
; Publication No. US20030143681A1
; GENERAL INFORMATION:
; APPLICANT: Immunex Corporation
; APPLICANT: Anderson, Dirk W.
; TITLE OF INVENTION: Human Ataxin-1-Like Polypeptide IMX97018
; FILE REFERENCE: 3138-A
; CURRENT APPLICATION NUMBER: US/10/207,706
; CURRENT FILING DATE: 2002-07-26
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 2
; LENGTH: 689
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-706-2

Query Match
Best Local Similarity 3.8%; Score 109.5; DB 14; Length 689;
Matches 120; Conservative 71; Mismatches 225; Indels 143; Gaps 30;

QY 5 SNDATPSNDG---AAGLVPSNNEAMALEPVPVQASLAAPVTGOTNIIDPWIRTNVQAPN 61
Db 33 STNHTPSSDASEWSRGVVWAGSQAGA-RVSLGDCGAEAITGLT--VDQGMLYKYAVPP 89
QY 62 GEFTVSPRSPGEILVNLE-LGPENL---PYLAHARMYN--GYAGMEVQVWMLAGNAFT 115
Db 90 ATF--SPTGLPS--VVMNSPLPPTFNVAASLIQHPGHHYPLHYAQLPSTSLQFISGPS 145
QY 116 AGKIIFAAVPPYPPFVENLSPS---QITMFPHVIDVRLTLEPVLPMDDVSTLPHENQKD 172
Db 146 ---LPYAVPPNPLPSPLLSPSANLATSHLPHFVFPYASLLAEGATPPPOAPSFAHSFNK-- 200
QY 173 EPKMLVAMLYTPLRSNGSGDDVFTVSCRILTRPSPEFDTYLVVPVTSKTKFTPLT--- 229
Db 201 -----APSATSPSGQ-----LPH--HSSTQPLDLAPG 225
QY 230 ---PVLTLGLBSNRPPLSIDEMVTS--NESIVQPPQNGRVTLDGEL-----LGT 275
Db 226 RMPYI-----YQMSRLPAGVTLTHTETPPAGASPVLTPOESQSALEAAANGGORPRERNLVR 281
QY 276 TQLQACNICSIRGKVTGQVPSQHMNLEITNLNGTQDPTDDVPAPL--GVPD----- 327
Db 282 RESEALDSPNSKGGGLVPVVECVVVGQL--FSGSQ--TPRVEVAAPAHRTGPTDPTDLEVQ 338
QY 328 -----PAGEVFGVLSQNRNGESNPANRAHDVAVATYDKYTPKGLVQIGTWTNDVENQ 382
Db 339 RVVGLAASQDYRVVAAQRKEEPSPLNLSH-----HTP-----DHQSGGR 377
QY 383 PTKFTPTIGLNEVANGHRF-----EQWTLPRYSGLTLMNMLAP--AVAPLFP-GER 430
Db 378 GSARNPAELAEKASQARGFYPSHQHEVFGHRPLPK--AMVANGNLVPTGTDGSLLPVSGE 435
QY 431 LLP-----FRSYVPLKGGFGNPAIDCSVPOEWHQHFQESAPSLGDLVALRVYVNPDTG 483
Db 436 ILVASSLDVQARATFPDKETPPPTTSSHL-----SHFMKGAIILQATGELKRVEDLQTO 491
QY 484 RVLFEAKLHKGGFLTVSST 502
Db 492 DFVRSAEV-SGGLKIDSST 509

```





```
QY 275 TTQLQACNICSRGKVTGQVPSEQHWNLE-----ITNLNGTQFDTDDVPAPLGVDPFA 329
Db 1044 TIVIPA-GSKSVSFIVDPTLEANQDAFNAEGMEKVVAITITGT-----SDNVTAATDIVDNA 1098
QY 330 GEV-----FGVLSQRNRGESNPANRAHDAV 354
Db 1099 GASATGVIIYDGNNAISLRNLGDFTLKYLSSSIAEKGFYITIGANGENDP-----MV 1152
QY 355 VATYSDKYPKLGVLQIGTWNT-----NDVENQP--TK-----FTPIGL----- 391
Db 1153 TTDYND--IVYGYYSQSKETTSYNSVANSQDNGPDGTTXDCNQSIITTVDLGAGDDLMI 1210
QY 392 --NEVAN-----GHRFEQWTLPRYSGALTNNMLAPAVAPLPFGERLLFFRSYVPLKGG- 443
Db 1211 RGNMLANTRVYTCGNNDTFTMDGMNTAL-----RVMYAGSYIFTESGD 1253
QY 444 -----FCNPAIDCS-VPQEWYQ 459
Db 1254 DIVTIKRTGVNAGQIYLGSGSDTFIQDATNNDTTLGGLDLGSGTKDISNMPKEYLS 1313
QY 460 HFYQESAPSLG-----DVALVRYNPDYGRVLFKAKLHKGGF---LTVSSTSTGP 506
Db 1314 VYQDGSNLSGNDNNIDTADYNTVYIYGSVSGEIL-----GGYGSNDNIIVTKNLTGN 1366
QY 507 VVVPANGYKFDSDWVNQFY--SLAPMGTCN 534
Db 1367 ISVGDNADTLTAGWI---YGGSTVSMGDCN 1393
```

Search completed: June 1, 2004, 14:04:40  
Job time : 37.5201 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.52349 Seconds  
(without alignments)  
4310.252 Million cell updates/sec

Title: US-09-926-799-7  
Perfect score: 2845  
Sequence: 1 MKMASNDATPSNDGAAGLVP.....VNQFYSLAPMTGNGRRRVQ 540

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1872.5	65.8	539	1 COAT LORDV	P54635 lordsdale v
2	1168.5	41.1	546	1 COAT_SOUV3	Q04542 southampton
3	321.5	11.3	2344	1 POLN_RHDV	P27410 rabbit hemo
4	318	11.2	702	1 COAT_SMSV1	P36284 san miguel
5	296.5	10.4	2208	1 POLN_MANCV	Q69014 manchester
6	292.5	10.3	668	1 COAT_FCVC6	P27404 feline cali
7	292.5	10.3	671	1 COAT_FCVF9	P27406 feline cali
8	280.5	9.9	668	1 COAT_FCVF4	P27405 feline cali
9	274	9.6	703	1 COAT_SMSV4	P36285 san miguel
10	143	5.0	2193	1 POLG_CX16T	Q94f31 c genome po
11	138	4.9	2193	1 POLG_CX16G	Q65900 c genome po
12	137	4.9	2333	1 POLG_FMDV1	P03306 f genome po
13	137	4.8	2332	1 POLG_FMDV0	P03305 f genome po
14	136	4.8	1011	1 POLG_FMDVT	P15072 foot-and-mo
15	131.5	4.6	2153	1 POLG_HRV16	Q82122 h genome po
16	129	4.5	2336	1 POLG_FMDV2	P49303 f genome po
17	127	4.5	2332	1 POLG_FMDVA	P03308 f genome po
18	124.5	4.4	2205	1 POLG_POL2W	P23069 p genome po
19	124	4.4	2194	1 POLG_EC30B	Q9wn78 e genome po
20	124	4.4	2194	1 POLG_HV701	P32537 h genome po
21	120.5	4.2	2207	1 POLG_POL2L	P06210 p genome po
22	118.5	4.2	2206	1 POLG_POL32	P06209 poliovirus
23	118	4.1	2208	1 POLH_POL1M	P03300 p genome po
24	117.5	4.1	5376	1 ZAN_MOUSE	O88799 mus musculu
25	116	4.1	2432	1 Y43R_IRV6	P18305 chilo iride
26	115.5	4.1	2175	1 POLG_BOVEV	P12915 b genome po
27	114.5	4.0	2184	1 POLG_EC01F	O91734 e genome po
28	114.5	4.0	2193	1 POLG_HE71B	Q66478 h genome po
29	114	4.0	2206	1 POLG_POL1M	P03299 p genome po
30	113	4.0	2185	1 POLG_SVDVU	P13900 s genome po
31	113	4.0	2193	1 POLG_HE71M	Q68479 h genome po
32	112.5	4.0	2206	1 POLG_POL3L	P03302 poliovirus
33	112	3.9	2209	1 POLG_POL1S	P03301 p genome po

34	110.5	3.9	521	1 VL2_HPV04	Q07862 human papil
35	110	3.9	832	1 POLG_HRV1A	P23008 human rhino
36	109	3.8	2183	1 POLG_CXB4E	Q86887 c genome po
37	108.5	3.8	1125	1 MAP4_MOUSE	P27546 mus musculu
38	108	3.8	2196	1 POLG_EC05N	Q9vlj1 e genome po
39	107.5	3.8	3828	1 TRX_DROVI	Q24742 drocephila
40	106.5	3.7	882	1 CADI_HUMAN	P12830 homo sapien
41	105.5	3.7	718	1 ATKX_AGR7S	Q8u9d9 agrobacteri
42	104.5	3.7	627	1 TESI_MOUSE	O70146 mus musculu
43	103.5	3.6	925	1 PIP1_YEAST	P40020 saccharomyc
44	103	3.6	519	1 ATL3_ARATH	Q9agh9 arabidopsie
45	103	3.6	2157	1 POLG_HRV1B	P12916 h genome po

ALIGNMENTS

RESULT 1					
ID	COAT LORDV	STANDARD;	PRT;	539 AA.	
AC	P54635;				
DT	01-OCT-1996 (Rel. 34, Created)				
DT	01-OCT-1996 (Rel. 34, Last sequence update)				
DT	01-NOV-1997 (Rel. 35, Last annotation update)				
DE	Coat protein (Capsid protein).				
OS	Lordsdale virus (Human enteric calicivirus).				
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;				
OC	Norovirus.				
OX	NCBI_TaxID=82658;				
RN	[1]_TaxID=82658;				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=96005060; PubMed=7561776;				
RA	Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;				
RT	"Human enteric Caliciviridae: the complete genome sequence and expression of virus-like particles from a genetic group II small round structured virus";				
RL	J. Gen. Virol. 76:2349-2355(1995).				
CC	- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.				
CC	THIS SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).				
CC	EMBL; X86557; CAA60255.1; InterPro; IPR004005; Calici_coat. InterPro; IPR008975; Viral_cap_coat. Pfam; PF00915; Calici_coat; 1. Coat protein; Glycoprotein. CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL). CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL). CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL). CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL). SEQUENCE 539 AA; 58775 MW; 585C63E7F2C5FD21 CRC64;				
DR	InterPro; IPR004005; Calici_coat.				
DR	InterPro; IPR008975; Viral_cap_coat.				
DR	Pfam; PF00915; Calici_coat; 1.				
KW	Coat protein; Glycoprotein.				
FT	CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).				
FT	CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).				
SQ	SEQUENCE 539 AA; 58775 MW; 585C63E7F2C5FD21 CRC64;				
Query Match	65.8%; Score 1872.5; DB 1; Length 539;				
Best Local Similarity	64.3%; Pred. No. 2.1e-130;				
Matches 348;	Conservative 83; Mismatches 103; Indels 7; Gaps 4;				
QY	1 MKMASNDATPSNDGAAGLVPVVGASLAAPVTGQNNIIDPWIRTNFVQAP 60				
Db	1 MKMASNDATPSNDGAAGLVPVVGASLAAPVTGQNNIIDPWIRTNFVQAP 60				
QY	61 NGEFTVSPRNPGEILVNLELGPENLYLAHARMYNGYAGMEVQVNLGNAGFTAGKII 120				
Db	61 GGEFTVSPRNPAGEILWSAPLGPLNLYSLHSLRMVNGYAGGFEVQVILGNAGFTAGKVI 120				
QY	121 FAAPPPPPFVENVSPSQITMFPVHVIDVRLLEPVLLEPMDVRLTFHFNQKDEPKRLVA 180				
Db	121 FAAPPPPPFTEGLSPSQVTMFPVHVIDVRLLEPVLLEPMDVRLTFHFNQKDEPKRLVA 180				



use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M57473; AAA47285.1; --  
 PIR; A41039; BRWHEB  
 PDB; 1KHV; 16-JAN-02.  
 MEROPS; C24.001; --

InterPro; IPR003593; AAA ATPase.  
 InterPro; IPR004005; Calici.coat.  
 InterPro; IPR004004; Calici\_pol.hel.  
 InterPro; IPR000317; Peptidase\_C24.  
 InterPro; IPR000605; RNA helicase.  
 InterPro; IPR007095; RNA pol DS.PS.  
 InterPro; IPR001205; RNA pol P3D.  
 InterPro; IPR007094; RNA pol\_PSVir.  
 InterPro; IPR008975; Viral\_cap.coat.  
 Pfam; PF00915; Calici.coat; 1.  
 Pfam; PF03510; Peptidase\_C24; 1.  
 Pfam; PF00680; RNA dep.RNA.pol; 1.  
 Pfam; PF00910; RNA helicase; 1.  
 PRINTS; PR00916; 2CENDOPTASE.  
 PRINTS; PR00918; CALICIVIRUSNS.

SMART; SM00382; AAA; 1.  
 Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;  
 Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.  
 KW CHAIN 7508 7707 HELICASE (P2C-LIKE).  
 FT CHAIN 1109 1251 PROTEASE P3C.  
 FT CHAIN 7197 71625 PROTEASE P3C.  
 FT CHAIN ? 2344 RNA-DIRECTED RNA POLYMERASE.  
 FT CHAIN ? 2344 COAT PROTEIN.  
 FT NP BIND 522 529 ATP (POTENTIAL).  
 FT ACT SITE 1212 1212 PROTEASE (POTENTIAL).  
 FT ACT SITE 1227 1227 PROTEASE (POTENTIAL).  
 SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 11.3%; Score 321.5; DB 1; Length 2344;  
 Best Local Similarity 25.9%; Pred. No. 3.1e-15;  
 Matches 152; Conservative 79; Mismatches 208; Indels 147; Gaps 30;

8 ATPNDGA-----AGLVPSNNEM-----ALPVPV-----GASLAAPVTG-----QTNIID 49  
 1772 AAPOGEAAGTATTASVPGTTDGMGVAATTSVITAENSSAIATAGIGPQQVDOQES 1831

50 PIRTNFVQAPNGEFTSPRNGPEILVNLELGPENLPYLAHLARMYNGYAGGMEVQVNL 109  
 1832 TW-RNFYY--NDVFTSWADAPGSLITYVQHSPPQNPPTAVLSQYAGWAGMGQFRIV 1888

110 AGNAFTAGKIIPAAVPPYPVENLSPS-QITMFPHVIIDVRLTPVLLPMPDVRSTLPHF 168  
 1889 AGSGVFGRLVRAVIPPGI---EIGFGLVRFQPHVVIDARSLEPVTITMPDLRPNMYH- 1944

169 NKQDEPKM--RLVAMLYTLRNSGGDDVFTVSCRILTRPSEFDTYLVPP---TVESK 223  
 1945 -PTGDFGLVFTLVSYNNL--INPFGSTSAIQVTETRPSEDFEFVIRAPSSKTVDSI 2002

224 TKP--FTLPLVTLGELNSRFLPSIDEMVTSNESIV-VOPQNGRT-----LDGELL 273  
 2003 SPAGLLTTPVLT--GVGNDNRW-----NGQIVGLQVPVGGFSTCNRHNNLNGSTY 2050

274 GTTQLOACNICIRGKVTQCPVSEQHMMNLEITNLNGTQFDPTDDVPAPLGVPDF----- 328  
 2051 GWSSPRFGDIDHRGASYSGSNATNLVQFWANAGSAIDNPISQV-APDGFPMDSFVPF 2109

329 -----AGEV--FGVLQRNRESNPAN--RAHDAVVAT-----YSDK 361  
 2110 NGFGIPAAAGWVGAIWNSGAPNVTTVQAYELGFATGACGNLOPNTTTSQAQTVAKSI 2169

362 YTPKGLVQ-----IGTWNVDVENQPT-----KFTPIGLN 392  
 2170 YAVVTGTAQNPAGLFVMAAGIISTENASAITTPOPDRIIVTFTGTAAPVGNKPTPIIPA 2229

393 EV-----ANCHRFQEWTLPRYSGALTNMNLAPAVAPLFPQERILLFRSYVP 439  
 2230 SVVTRTGDVNATAGSANGTQGTGSQLP---VTTLGLSLNNYSALMPGQ---FFVWOLT 2283  
 440 LKGGFGNPAIDCSVPQEWQHFYQESAPS-----LGDVALVRYNVP 480  
 2284 PASGP--MEIGLSVD---GYFYAGTGASTLIDLTTELIDVRPVP 2323

RESULT 4  
 COAT\_SMSV1

ID\_COAT\_SMSV1 STANDARD; PRT; 702 AA.  
 AC P36284;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DE 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS San Miguel sea lion virus (serotype 1) (SMSV 1).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OX NCBI\_TaxID=36406;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92410750; PubMed=1529644;  
 RA Neill J.D.;

"Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus: identification of conserved and non-conserved amino acid sequences among calicivirus capsid proteins."; Virus Res. 24:211-222(1992).  
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announcement/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; M87481; AAA16217.1; --  
 PIR; A48562; A48562.  
 InterPro; IPR004005; Calici.coat.  
 InterPro; IPR008975; Viral\_cap.coat.  
 Pfam; PF00915; Calici.coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 493 493 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 545 545 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DE3D7 CRC64;

Query Match 11.2%; Score 318; DB 1; Length 702;  
 Best Local Similarity 25.8%; Pred. No. 1e-15;  
 Matches 153; Conservative 74; Mismatches 199; Indels 168; Gaps 33;

8 ATPNDGAAGLVPSNNEMALEPVMGAS-----LAAPVTGQTNIID-PWIRTNFVQ---A 59  
 151 AESDPGCGADIVTEBQGTVVQQPVPQAQSLTTAAASTGKT--VDCW--TTFFSYHTA 206

60 PNGEFTVSPRNSPGBILVNLELGPENLPYLAHLARMYNGYAGGMEVQVNLNAGNATAGKI 119  
 207 VNMSTT---BAQKGLFSLRALSPENLPYLRHISLSYSTWSGGIDVRETVSGSGVFGKL 262

120 IFAAVPPYP-PVENLSPSQTITMFPHVIIDVRLTPVLLPMPDVRSTLPHFNQKDPKML 178  
 263 AALIVPPGIEPVE--SPIML-QYPHVLPDARQTEPVTITPIDIKTLTH-SMDDTDTTL 318

179 VAMLYTPI-RLNSGGDDVFTVSCRILTRPSEFDTYLVPPPTVBSKTKPFTLPVLTLGEL 237  
 319 VIMVYNELINPVEQSEPKSSCSITVETRPSSDFTFSLKKPPG-----SLKKGSI 368

238 SNRPPLSIDEMVTSNESIVVQPNQR-----VTLDGELLTTLQACNICIRK 289

369 -----PSDLI---PRNSRHMGNRWMSTIDGFFVQPRVFSQNRHFDST 410  
Db POLN MANCV STANDARD; PRT; 2208 AA.  
AC O69014;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polyprotein [Contains: RNA-directed RNA polymerase  
DE (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like  
DE protein); Coat protein].  
OS Manchester virus (human enteric calicivirus).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Sapovirus.  
OC NCBI\_TaxID=82659;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95390791; PubMed=7661689;  
RA Liu B.L., Clarke J.N., Caul E.O., Lambden P.R.;  
RT "Human enteric caliciviruses have a unique genome structure and are  
RT distinct from the Norwalk-like viruses."  
RL Arch. Virol. 140:1345-1356(1995).  
CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).  
CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR  
CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY  
CC SIMILARITY)  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA} (N).  
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; X86560; CAA60262.1; --  
DR MEROPS; C24.UPW; --  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR004004; Calici\_pol\_hel.  
DR InterPro; IPR000317; Peptidase\_C24.  
DR InterPro; IPR000605; RNA\_helicase.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR InterPro; IPR007094; RNA\_pol\_P5Vir.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.

DR PF03510; Peptidase\_C24; 1.  
DR Pfam; PF00860; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
DR PRINTS; PR00916; 2CENDOPTASE.  
DR PRINTS; PR00918; CALICIVIRUSNS.  
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;  
KW Thiol protease; Helicase; ATP-binding; Coat protein.  
FT CHAIN ? ? HELICASE (P2C-LIKE).  
FT CHAIN ? ? PROTEASE (P3C).  
FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.  
FT CHAIN ? ? COAT PROTEIN.  
FT NP\_BIND 408 415 ATP (POTENTIAL).  
FT ACT\_SITE 1097 1097 PROTEASE (BY SIMILARITY).  
FT ACT\_SITE 1112 1112 PROTEASE (BY SIMILARITY).  
FT ACT\_SITE 1112 1112 PROTEASE (BY SIMILARITY).  
SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;  
Query Match 10.4%; Score 296.5; DB 1; Length 2208;  
Best Local Similarity 22.4%; Pred. No. 2e-13;  
Matches 144; Conservative 77; Mismatches 231; Indels 191; Gaps 28;  
QY 2 KMASNDATPSNDGAGLVPESSNEAMALBPVVGASLAAPVTGQTNIIDPWIRTNFVOAPN 61  
Db 1647 EMEGNSNP-----PKQSNPNVVDP---PGTTGPTTSHVVANP-----EQPN 1688  
QY 62 G-----EFTVSPRNSPGEILNVLNLELGPENLYLAH 91  
Db 1689 GAAQRLAVATGAIQSNVPEAIRNCFVPRTFAMNDRMPTGTFLGSLILHPNINPTSH 1748  
QY 92 LARMYNGYAGMEVQVWMLAGNAFTAGKLIIFAAVPPYFPVENLSPSQIT---MPPHVIIDV 148  
Db 1749 LSGWAGWGGSFVRLSIGSGVFGAGRIIASVIPP-----GVDPSLRDPGLVPHAFVDA 1803  
QY 149 RTLPVLLPMDVRSITLHFNQDEPKQRLVAMLYTPLRNSGSGDDVFTVSCILIRPSP 208  
Db 1804 RITEPVFMPIDVRAVDYHRMDGAEPTCSLGFVYQPLNPPFTAVSTCWVSEVETKPGG 1863  
QY 209 EFDFTYLVPP--TVESKTKFTPLVLTGELSNSRFPFLSIDEMVTSNESISVVPQNGRV 266  
Db 1864 DFDFCLLRPPGQOMENGVSPEGLLPRRLGYSRGRNVGGLVGVGLLVAEHKQVNRHFNNS 1923  
QY 267 TLDGELLGTTQLQACNICSIRKVTGQVPS--EQHMNLEITNLNGTQDPTDDVPAPLGV 325  
Db 1924 VTFGWSHTAPVNPMAEI-----VTNQAHSSTRHAW--LSIGAQNKGFLP-----GI 1968  
QY 326 PDF-----AGRVFGLV-----SQNRGESNPA-----NRAHDVVATV----- 358  
Db 1969 PNHFPDSCASTVVGAMDTSLGRFPSTGCGPAISFQNNQGVYENDTFSVMFATYDPLTSG 2028  
QY 359 -----SDKYTP-KLGLVOIG-----TW-----NTNDVENQPT----- 384  
Db 2029 TGVALTNSINPASLALVRISNNDFTSGFANDKNVWQMSWEMVTGNIQIRGQVTPMSGT 2088  
QY 385 --KFTPIGLNEVANGHRFEQMTLPRYSGALTINMNLAPAVAPLPFGERLLFFRSVYVPLKG 442  
Db 2089 NYTFTSTGANTLV-----LMQ-----ERMLSYD----- 2111  
QY 443 GFGNPAIDCSVQV 502  
Db 2112 --GHQAIIYSSQLERTAEYFQNDIVNIIPENSMAVF--NVETHSASFQIGRDPGYM-VITGG 2167  
QY 503 STGPVVVPVANGYFKFDSWVQVYSLAPMG-----TGN-GRRR 538  
Db 2168 SIG-VNVFLEPETRF-----QYVGILPLSALSGPSGNMGRAR 2204

RESULT 6  
COAT\_FCV66  
ID\_COAT\_FCV66 STANDARD; PRT; 668 AA.  
AC P27404;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Coat protein (capsid protein).

OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
 OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OX NCBI\_TaxID=11979;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=91374597; PubMed=1716692;  
 RA Neill J.D., Reardon I.M., Heinrikson R.L.;  
 RT "Nucleotide sequence and expression of the capsid protein gene of  
 RT feline calicivirus";  
 RL J. Virol. 65:5440-5447 (1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Neill J.D.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- PTM: The N-terminus is blocked.  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M32819; AAA42925.1; -;  
 DR EMBL; U13992; AAC13993.1; -;  
 DR PIR; A40507; VCMWFF.  
 DR InterPro; IPR004005; Calici.coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 668 AA; 73550 MW; 9E52312108D4D41 CRC64;  
 Query Match 10.3%; Score 292.5; DB 1; Length 668;  
 Best Local Similarity 26.0%; Pred. No. 7.3e-14;  
 Matches 113; Conservative 53; Mismatches 143; Indels 125; Gaps 21;  
 QY 11 SNDGAAGLVPSNN--EAMALEPVVGASLAAPVTGQTNIIDPW-IRTNFVOAPNGEFTVS 67  
 DB 125 ADDGSI-TTPEQGTWGGVIAEPNQMSTADMATKGSVDSEWEAFFSFHTSVNW----S 179  
 QY 68 PRNSPGEILVNLELGPPELNPYLAHLMRYNGYAGGMEVQVLMAGNAFTAGKIIFAAVPPY 127  
 DB 180 TSETQKILFKQSLGFLNPLNLYTHLAKLYVAMSGSDVRFSGSGVFGGKLAIVVPPG 239  
 QY 128 P-PVENLSPSQITMPEPHVVDVRLTLEPVLPMVDVRSITLPHFNQKDEPKMRVLAIVTPL 186  
 DB 240 IDPVQSTSMQ---YPHFLDARQVEPVIFSPIDRLSTLYHL-MSDPTTTSILVIMVNDL 295  
 QY 187 -----RNSGDDVFTVSC--RILTRPSPEFDFTYLVPPVSKTKPFTLPVLTGLSLS 240  
 DB 296 INPYANDNSSG-CIVTVE-----TKPGDQFKHLKPPG-----SMLTGSI--- 337  
 QY 241 RPPLSIDEMVTSPNESIVVQPQNGRVTLDGELLGTTQLQACNICISIRGKVTGQVPSQHM 300  
 DB 338 -----PSDLI-----PKSSSL 348  
 QY 301 WNLTEITNLNGTQF--DPTDDVPAPLGVDPDFPAGEVGV-----LSQRNGESNPANRAHDA 353  
 DB 349 W-----IGNRFWDITDFVIRPF-----VFQANRHFDFNQETAGWSTPRP---P 390  
 QY 354 VVATYSDKYTPKLGVLQVIGT-----WNTNDVENQPTKFTPIGLNEVANGHFRFQW 403  
 DB 391 ITITISVESAKLG-IGVATDVIYVIGIPDGWPDPTTI-----PGELVPVGYAITNGTNDIT 446

QY 404 TLPFRYSGALTLNMN 417  
 DB 447 TAAQYDAATEIRNN 460  
 RESULT 7  
 COAT\_FCVF9 STANDARD; PRT; 671 AA.  
 ID COAT\_FCVF9 AC P27406;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Coat protein (Capsid protein).  
 GN CYP76.  
 OS Feline calicivirus (strain F9) (FCV).  
 OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OX NCBI\_TaxID=11981;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92410623; PubMed=1529544;  
 RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,  
 RA Turner P.C.;  
 RT "The complete nucleotide sequence of a feline calicivirus.";  
 RL Virology 190:443-448 (1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92339673; PubMed=1633955;  
 RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;  
 RT "Cloning and sequence determination of the feline calicivirus strain  
 RT F9.";  
 RL Biochem. Soc. Trans. 20:265-268 (1992).  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; M86379; AAA79327.1; -;  
 DR EMBL; Z11536; CAA77636.1; -;  
 DR PIR; B43382; VCMWF9.  
 DR InterPro; IPR004005; Calici.coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 671 AA; 73441 MW; 33BE86D8370D5ES CRC64;  
 Query Match 10.3%; Score 292.5; DB 1; Length 671;  
 Best Local Similarity 22.7%; Pred. No. 7.3e-14;  
 Matches 133; Conservative 84; Mismatches 223; Indels 147; Gaps 25;  
 QY 11 SNDGAAGLVPSNN--EAMALEPVVGASLAAPVTGQTNIIDPW-IRTNFVOAPNGEFTVS 67  
 DB 125 ADDGSI-TAPEQGTWGGVIAEPNQMSTADMATKGSVDSEWEAFFSFHTSVNW----S 179  
 QY 68 PRNSPGEILVNLELGPPELNPYLAHLMRYNGYAGGMEVQVLMAGNAFTAGKIIFAAVPPY 127  
 DB 180 TSETQKILFKQSLGFLNPLNLYTHLAKLYVAMSGSIEVRFSGSGVFGGKLAIVVPP- 238  
 QY 128 PVENLSPSQITMPEPHVVDVRLTLEPVLPMVDVRSITLPHFNQKDEPKMRVLAIVTPLR 187  
 DB 239 -GVDPVQSTSMQTPHVLFDARQVEPVIFCPLDRLSTLYHL-MSDPTTTSILVIMVNDL- 295  
 QY 188 SNGSGDDVFTVSC--RILTRPSPEFDFTYLVPP-----TVESKTKPFTLPVLTGLSLS 238

b	296	INPYANDANSSGCI	VTVETKPGDPFKFHLKPPGSM	LTHGSI	PSDLIPKT	SSLWIG	---	351	
y	239	NSRFPLSIDEM	TSP	---	NE	---	SIVVQPONGR	265	
b	352	NRYSMDITDF	VIRPFVFOANRHF	NOETAG	SWTPRFRP	ISVITTEQ	NGAKLIGVATD	410	
y	266	---	VTLDGELLGTTQ	LAQC	---	ICSIRG	KVTGOV	295	
b	411	YIVPGIPDGP	DDTTPIGELI	PAGDYAIT	NGTNDIT	TATGYTAD	IIKNNINFRGM	470	
y	296	SEOHMMNLEI	TNLNGTQF	DDVPAPL	GVPDFAGE	FCVL	SQRNGESNPANRAHDAV	355	
b	471	SLQRAMGDK	-KISNTAFIT	---	ATLGD	NNKINPCNT	IDQSKI	511	
y	356	ATYSKDYKTK	GLVQIGTWN	TNDVENQ	PTKFTPI	IGLNEVANGH	RFEQ	411	
b	512	VVFQDNHVGK	---	QTSDDT	LALLGYT	-GIGEQA	IGSDRDRVVRIS	TLP	561
y	412	LTLNMLAPAV	APLFFGER	LLPFRS	VVPLKGGFN	PAIDCSVP	QEWVHFQYESAP	SLGD	471
b	562	R	---	CGNHPIFYK	SIKL	-GVVIRS	IDV	---	596
y	472	VALRVYNPD	TGTVLPEAKLHKG	FLTVS	TSSTG	PPVVPANG	YKFPD	518	
b	597	LSLNHYLLP	DDPSFAVYRI	DSNGSW	PDIGID	SGDFSV	GVSGFGKLE	643	

RESULT 8

COAT_FCVF4	STANDARD;	PRT;	668 AA.
AC P274051			
AD 01-AUG-1992 (Rel. 23, Created)			
AE 01-AUG-1992 (Rel. 23, Last sequence update)			
AF 01-JUN-1994 (Rel. 29, Last annotation update)			
AG Coat protein (Capsid protein)			
AA Feline calicivirus (strain Japanese F4) (FCV)			
AC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae; Vesivirus.			
AD NCBI_TaxID=11980;			
AE [1]			
AF SEQUENCE FROM N.A.			
AG MEDLINE=91306470; PubMed=1853578;			
AA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N., Miyamura K., Yamazaki S., Mikami T.,			
AT "Sequence analysis of the 3'-end of feline calicivirus genome."			
AR Virology 183:810-814 (1991).			
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.			
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announcement">http://www.isb-sib.ch/announcement</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC EMBL; D90357; BAAL4371.1; -			
DR PIR; B40481; VCVWFC.			
DR InterPro; IPR004005; Calici_coat.			
DR InterPro; IPR008975; Viral_Cap_coat.			
DR Pfam; PF00915; Calici_coat; 1.			
DR Coat protein; Glycoprotein.			
KW CARBOHYD 177 177			
FT N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT CARBOHYD 301 301			
FT N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT CARBOHYD 304 304			
FT N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT CARBOHYD 399 399			
FT N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT CARBOHYD 459 459			
FT N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT CARBOHYD 615 615			
FT N-LINKED (GLCNAC. . .) (POTENTIAL).			
FT SEQUENCE 668 AA; 73589 MW; 85BDBC85804E503 CRC64;			

Query Match 9.9%; Score 280.5; DB 1; Length 668;  
Best Local Similarity 32.0%; Pred. No. 5.6e-13;



Db 672 VNWDIDLNGYALRRKCLFTYMRDAEFTFVAKNGELVPQLLQYVVPQAPKPTSR 731  
QY 301 -----WNLEITNLNGTQFDPTDDVPAPLQVP-----DFA 329  
Db 732 DSFAHQ---TATNPSVFKVMTDPPAQVSPFMSPASAYQWFDGYPTFGHQLANDLDYG 788  
QY 330 -----GEVFGVLSORNG-ESNPANRAHDAVATYSKYTPKLGVLQIGTWTNDVENOPTK 385  
Db 789 QCPNNMGTFSIRTGVTEKSP-----HSTLRVY-----WRIKHVRAWIIPRLRNQPYL 837  
QY 386 FTPIGLNEVANGHREFTQWTLPRYS 410  
Db 838 FK-----INPNYKG 846

RESULT 11

POLG CX16G  
ID -POLG CX16G STANDARD; PRT; 2193 AA.  
AC Q65900;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 10-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].  
OS Coxsackievirus A16 (strain G-10).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.  
OC NCBI\_TaxID=69159;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=94303216; PubMed=8030260;  
RA Poyry T., Hyypiae T., Horsnell C., Kinnunen L., Hovi T., Stanway G.;  
RT "Molecular analysis of coxsackievirus A16 reveals a new genetic group of enteroviruses.";  
RL Virology 202:982-987(1994).  
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular membranes and is associated with viral RNA synthesis.  
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA) (N).  
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.  
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U05876; AAA50478.1; --  
CC HSSP; P03300; IPOV.  
CC MEROPS; C03.022; --  
CC MEROPS; C03.022; --  
CC InterPro; IPR003593; AAA\_ATPase.  
CC InterPro; IPR004004; Calici\_pol\_hel.  
CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
CC InterPro; IPR000199; Pept\_3C\_picorn.  
CC InterPro; IPR000081; Peptidase\_C3.  
CC InterPro; IPR003136; Pico\_P1A.  
CC InterPro; IPR002527; Pico\_P2B.  
CC InterPro; IPR001676; Rhv.

DR InterPro; IPR000605; RNA\_helicase.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00548; Cys-protease-3C; 1.  
DR Pfam; PF02226; Pico\_P1A; 1.  
DR Pfam; PF00947; Pico\_P2A; 1.  
DR Pfam; PF01552; Pico\_P2B; 1.  
DR Pfam; PF00073; rhv; 3.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
DR PRINTS; PR00918; CALICVIRUSNS.  
DR ProDom; PD001125; Cys\_protease\_3C; 1.  
DR ProDom; PD001306; Pico\_P2A; 1.  
DR ProDom; PD001274; Pico\_P2B; 1.  
DR SMART; SM00382; AAA; 1.  
KW Polyprotein; Coat protein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
KW Lipoprotein.  
FT CHAIN 2 69 COAT PROTEIN VP4.  
FT CHAIN 70 323 COAT PROTEIN VP2.  
FT CHAIN 324 565 COAT PROTEIN VP3.  
FT CHAIN 566 862 COAT PROTEIN VP1.  
FT CHAIN 863 1012 CORE PROTEIN P2A.  
FT CHAIN 1013 1111 CORE PROTEIN P2B.  
FT CHAIN 1112 1440 CORE PROTEIN P2C.  
FT CHAIN 1441 1526 CORE PROTEIN P3A.  
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VP3.  
FT CHAIN 1549 1731 PICORNAIN 3C.  
FT CHAIN 1732 2193 N-myristoyl glycine (in host) (By similarity) (POTENTIAL).  
FT ACT SITE 1695 1695 PROTEASE (POTENTIAL).  
FT ACT SITE 1709 1709 PROTEASE (POTENTIAL).  
SQ SEQUENCE 2193 AA; 243209 MW; 04B3BCE572A76E38 CRC64;  
Query Match 4.9%; Score 138; DB 1; Length 2193;  
Best Local Similarity 20.5%; Pred. No. 0.094;  
Matches 101; Conservative 53; Mismatches 185; Indels 154; Gaps 21;  
QY 44 OTNIIDPWIRNF---VQAPNGEFTVSPRNSPGBILVNLELGPENLPYLAHLARMYNGYA 100  
Db 382 QSNETTFMQLCFPVSVQSKTGELCAVFRADPG-----RNGPMQSTILGOLCRYTQWS 435  
QY 101 GMEVQVVMAGNAFTAGKIIFAANVPP--YFVENLSPSQITMFPHPHVIDVRLTLEPVLPM 158  
Db 436 GSLEVTFMFAGSFMATGKMLIAYTPPGGVPADRLTAMLGT---HVIWDFGLQSSVTLVI 492  
QY 159 PDVRSTLFHFNKD-----EPKRLVAML-----YT----- 184  
Db 493 PWISNTHYPAHAKDGYFDYTTGTITIWYQYVVPVIGAPTTAYIVALAAQDNFTMKLC 552  
QY 185 -----PLRNSGSDVFTVSCRILTRPSPDFTYLVPPTV-----ESTKPTFLPVL 232  
Db 553 KOTEDIEQSANIQGDGIADMDQAVTSRVGRALTSQVEPTAANTNASEHRLGTGLVPAL 612  
QY 233 TLGEL---SNSRFFLSID-----EMVTSFN-----ESTVQPNQGRVTLDG-- 270  
Db 613 QAAETGASSNAQENLLETRCVLNHHSIQETTIGNFFSRAGLVSITMPTTGTQNTDGYV 672  
QY 271 -----ELIGTTQL-QACNICS-----IRKVTQGVPSF---QHM----- 300  
Db 673 NWDIDLNGYAQMRRKCELFYMRDAEFTFVAKNGELVPQLLQYVVPQAPKPTSRD 732  
QY 301 -----WNLEITNLNGTQFDPTDDVPAPLQVPDFAG-----EVFGVLSORNGESN- 345  
Db 733 SFAMQ---TATNPSIFVKLTDPPAQVSPFMSPASAYQWFDGYPTFGAHPQSDNADYQ 789  
QY 346 -PANRAHDAVATYSKYTPK-----LGLVOIGTWTNDVENOPTKFTPIGLNEVANG 397  
Db 790 CPNNMGTFSIRTGVTEKSPHSITLRVYRIKHVRAWIIPRLRNQPYLFK----- 839





FT DISULFID 511 INTERCHAIN (IN VP3 DIMER).  
 FT DISULFID 406 IN VP2-VP1 DIMER.  
 FT VARIANT 780 I -> V (IN STRAIN O1BFS).  
 FT VARIANT 808 G -> R (IN STRAIN O1BFS).  
 FT VARIANT 861 N -> S (IN STRAIN O1BFS).  
 SQ SEQUENCE 2332 AA; 258924 MW; 4A83176F43447D68 CRC64;  
 Query Match 4.8%; Score 137; DB 1; Length 2332;  
 Best Local Similarity 21.6%; Pred. No. 0.12;  
 Matches 77; Conservative 40; Mismatches 130; Indels 110; Gaps 14;  
 QY 11 SNDGAAGLVP-----ESNNE-----AMALEPVCASLAAPVTGOTNIIDPWIRTFVQA 59  
 DB 248 SNEGSDTTTTHNTQNNDFSKLASSAFGLFALLADKKTETTLLEDRILT-----T 303  
 QY 60 PNCGETVSPNSPG-----ELLVN----- 78  
 DB 304 RNGTHSTTQSSVGVTVGYATADDFVSGNTSGLETRVVQAERFFKTHLFDWVTSDFGR 363  
 QY 79 ---LELGPENLPYLAHLARMYAGYAGMEVQVNLGNAFTAGKIIPAAVPPFPVENLSP 135  
 DB 364 CHLELPDTHKGVYGLTDSYAMRNGWDVETAVGNQFNGGCLLVAMVPELYSIQKREL 423  
 QY 136 SQITMPHVIDVRLTPVLLPMPDYRSTLPHNQKDEPKMR-----LVAMLYTPLRNSGS 191  
 DB 424 YQLTLFPHQFINPRTNMTAHTVTFV-----GVNRYDQYKVKPWTLLVVMVAPLTVNTE 478  
 QY 192 GDDVFTVSCRLTRPSPDFPTLVLPVTVESKTKPFTPLVTLGELNSNR--PPLSIDE- 248  
 DB 479 G-----APQIKVYANIAPT-----NVHVAGEFPSEKGIFFVACSDG 514  
 QY 249 ---MVTSPNESIVVQNGRV-----TLGELLGTTQLQACNICSIRGKVTGQVP 295  
 DB 515 YGLVTTDPK---TADPVYGVKFNPPRNQLPGRF--TNLLDVAEACFTFLRFEGGVP 566  
 RESULT 14  
 POLG FMDVT STANDARD; PRT; 1011 AA.  
 ID POLG FMDVT AC P15072; Q84755; Q84757; Q84758;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-NOV-1990 (Rel. 16, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat  
 DE protein VP4; Coat protein VP2; Coat protein VP3; Coat protein VP1;  
 DE Core protein P12] (Fragment).  
 OS Foot-and-mouth disease virus (strain C1) (Aphthovirus C) (FMDV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Aphthovirus.  
 OC NCBI\_TaxID=12121;  
 RN [1]  
 RP MEDLINE=84069809; PubMed=6316275;  
 RA Beck E., Forss S., Strebel K., Cattaneo R., Feil G.;  
 RT "Structure of the FMDV translation initiation site and of the  
 RT structural proteins";  
 RL Nucleic Acids Res. 11:7873-7885(1983).  
 CC each of which is composed of 60 icosahedral units,  
 CC VP3, and VP4.  
 CC -1- SUBUNIT: The virus capsid is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -1- MISCELLANEOUS: THE COAT PROTEIN VP1 CONTAINS THE MAIN ANTIGENIC  
 CC DETERMINANTS OF THE VIRION; THEREFORE, CHANGES IN ITS SEQUENCE  
 CC MUST BE RESPONSIBLE FOR THE HIGH ANTIGENIC VARIABILITY OF THE  
 CC VIRUS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

RL Nucleic Acids Res. 12:6587-6601(1984).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O1BFS;  
 RX MEDLINE=83143292; PubMed=6298715;  
 RA Makoff A.J., Paynter C.A., Rowlands D.J., Boothroyd J.C.;  
 RT "Comparison of the amino acid sequence of the major immunogen from  
 RT three serotypes of foot and mouth disease virus";  
 RL Nucleic Acids Res. 10:8285-8295(1982).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RX MEDLINE=89143740; PubMed=2537470;  
 RA Acharya R., Fry E., Stuart D., Fox G., Rowlands D., Brown F.;  
 RT "The three-dimensional structure of foot-and-mouth disease virus at  
 RT 2.9-A resolution.";  
 RL Nature 337:709-716(1989).  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA} (N).  
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -1- MISCELLANEOUS: THE STRAIN O1K SEQUENCE IS SHOWN  
 CC -1- MISCELLANEOUS: THE COAT PROTEIN VP1 CONTAINS THE MAIN ANTIGENIC  
 CC DETERMINANTS OF THE VIRION; THEREFORE, CHANGES IN ITS SEQUENCE  
 CC MUST BE RESPONSIBLE FOR THE HIGH ANTIGENIC VARIABILITY OF THE  
 CC VIRUS.  
 CC -----  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; X00871; CAA25416.1; -;  
 CC EMBL; J02185; AAA42635.1; -;  
 CC PDB; 1QMY; 18-SEP-01.  
 CC PDB; 1QOL; 10-NOV-00.  
 CC MEROPS; C03.008; -;  
 CC InterPro; IPR004004; Calici\_pol\_hel.  
 CC InterPro; IPR009003; Cys\_Ser\_trypsin.  
 CC InterPro; IPR004080; FMDVPicoat.  
 CC InterPro; IPR008739; Peptidase\_C28.  
 CC InterPro; IPR001676; Rhv.  
 CC InterPro; IPR000605; RNA helicase.  
 CC InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 CC InterPro; IPR001205; RNA\_pol\_P3D.  
 CC InterPro; IPR007094; RNA\_pol\_Psvir.  
 CC InterPro; IPR008975; Viral\_cap\_coat.  
 CC Pfam; PF05408; Peptidase\_C28; 1.  
 CC Pfam; PF00073; rhv; 3.  
 CC Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 CC Pfam; PF00910; RNA\_helicase; 1.  
 CC PRINTS; PR00918; CALICVIRUSNS.  
 CC PRINTS; PR01542; FMDVPICOAT.  
 CC PolyProtein; Coat protein; Core protein; RNA-directed RNA polymerase;  
 CC Transferase; Hydrolase; Thiol protease; Nonstructural protein;  
 CC Myristate; 3D-structure; Lipoprotein.  
 FT CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.  
 FT CHAIN 202 286 COAT PROTEIN VP4.  
 FT CHAIN 287 504 COAT PROTEIN VP2.  
 FT CHAIN 505 724 COAT PROTEIN VP3.  
 FT CHAIN 725 937 COAT PROTEIN VP1.  
 FT CHAIN 938 1107 CORE PROTEIN P12.  
 FT CHAIN 1108 1425 CORE PROTEIN P34.  
 FT CHAIN 1426 1578 CORE PROTEIN P14.  
 FT CHAIN 1579 1643 GENOME-LINKED PROTEIN VP6.  
 FT CHAIN 1650 1862 PROTEASE.  
 FT CHAIN 1863 2332 RNA-DIRECTED RNA POLYMERASE.  
 FT CHAIN 202 N-myristoyl glycine (in host).  
 FT LIPID



DR Pfam; PF00910; RNA\_helicase; 1.  
DR PRINTS; PR00918; CALICVIRUSNS.  
DR ProDom; PD001125; Cys\_protease\_3C; 1.  
DR ProDom; PD001306; Pico\_P2A; 1.  
DR ProDom; PD001274; Pico\_P2B; 1.  
DR SMART; SM00382; RAA; 1\_  
KW Polyprotein; Coat protein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
KW 3D-structure; Lipoprotein.  
FT CHAIN 2 69  
FT CHAIN 70 330 COAT PROTEIN VP4.  
FT CHAIN 331 568 COAT PROTEIN VP2.  
FT CHAIN 569 853 COAT PROTEIN VP3.  
FT CHAIN 854 985 COAT PROTEIN VP1.  
FT CHAIN 986 1090 CORE PROTEIN P2A.  
FT CHAIN 1091 1412 CORE PROTEIN P2B.  
FT CHAIN 1413 1489 CORE PROTEIN P2C.  
FT CHAIN 1490 1510 CORE PROTEIN P3A.  
FT CHAIN 1511 1693 GENOME-LINKED PROTEIN VPG.  
FT CHAIN 1694 2153 PICORNAIN 3C.  
FT LIPID 2 2  
FT ACT\_SITE 1657 1657 RNA-DIRECTED RNA POLYMERASE P3D.  
FT ACT\_SITE 1671 1671 N-myristoyl glycine (in host).  
FT CONFLICT 547 548 PROTEASE (POTENTIAL).  
FT STRAND 27 30 KD -> NH (IN REF. 1).  
FT HELIX 36 38  
FT TURN 80 81  
FT STRAND 83 87  
FT TURN 88 89  
FT STRAND 90 94  
FT STRAND 101 102  
FT HELIX 103 105  
FT TURN 113 115  
FT STRAND 123 123  
FT TURN 126 128  
FT TURN 129 129  
FT STRAND 133 134  
FT STRAND 138 140  
FT TURN 142 143  
FT STRAND 147 151  
FT TURN 152 152  
FT HELIX 153 155  
FT TURN 156 157  
FT STRAND 159 167  
FT STRAND 168 180  
FT TURN 185 186  
FT STRAND 188 197  
FT TURN 198 198  
FT STRAND 204 204  
FT TURN 207 208  
FT HELIX 213 216  
FT STRAND 219 221  
FT STRAND 223 224  
FT TURN 225 227  
FT TURN 232 233  
FT HELIX 239 241  
FT TURN 242 245  
FT STRAND 246 246  
FT HELIX 248 253  
FT STRAND 256 260  
FT TURN 261 263  
FT STRAND 266 271  
FT TURN 280 280  
FT TURN 282 284  
FT STRAND 285 285  
FT STRAND 288 288  
FT TURN 303 304  
FT STRAND 307 323  
FT TURN 338 341  
FT TURN 345 346  
FT STRAND 353 353  
FT TURN 356 357  
FT STRAND 369 370

FT STRAND 372 372  
FT TURN 373 373  
FT HELIX 374 377  
FT TURN 378 378  
FT STRAND 381 382  
FT TURN 385 386  
FT STRAND 388 388  
FT TURN 389 393  
FT HELIX 398 398  
FT STRAND 399 402  
FT TURN 408 409  
FT STRAND 411 416  
FT TURN 419 420  
FT TURN 422 426  
FT TURN 428 433  
FT TURN 434 435  
FT STRAND 436 440  
FT STRAND 443 449  
FT TURN 453 454  
FT STRAND 456 456  
FT STRAND 458 464  
FT HELIX 474 478  
FT TURN 479 479  
FT STRAND 481 486  
FT STRAND 492 497  
FT STRAND 506 507  
FT TURN 513 514  
FT STRAND 518 523  
FT STRAND 528 528  
FT TURN 531 532  
FT STRAND 537 545  
FT TURN 547 548  
FT STRAND 550 554  
FT STRAND 567 567  
FT HELIX 570 579  
FT TURN 580 580  
FT STRAND 584 584  
FT STRAND 588 588  
FT TURN 600 601  
FT STRAND 602 603  
FT HELIX 605 607  
FT HELIX 615 618  
FT STRAND 624 624  
FT STRAND 629 629

Query Match 4.6%; Score 131.5; DB 1; Length 2153;  
Best Local Similarity 20.8%; Pred. No. 0.28; Indels 173; Gaps 24;  
Matches 108; Conservative 61; Mismatches 177;  
QY 48 IDPWIRTNFVQAPNGE-----FTVSPRNSPGE-----ILVNLELGPENLPVLAHLARMYNG 98  
Db 379 VDTLIPINSTQSNIGNVSMYTVTLSPQTKLAEIEFAIKVDIASHPATLTLIGEIASYFTH 438  
QY 99 YAGGMEVQVNLGNAFTAGKIIFAAVPPYFPVENLSPSQITMFPHVIVDTLPVLLPM 158  
Db 439 WTGSLRFPFPMFCGTANTTLKVLAVTPPGIG-KPRSRKEAMLGTHVVMVDVGLQSTVSLVV 497  
QY 159 PDVRSTLPHFNQKDEPKMLVAMLYTPLRNSGSDVFTVSCRILTRPSPEFDFTYLVPP 218  
Db 498 PWISASQYRFT-----TPDTYSAG-----YITCWYQT-----NFVVPP 531  
QY 219 -----TVE-----SKTKPFTLPVLTIGELSNRPPLS-----IDEMVTSNESIVV--- 259  
Db 532 NTENTAEMLCFVSGCKDFCLRMARDTLHKQTGPITQNPVERVYDEVL---NEVLVWPN 588  
QY 260 -----QPQNGRVTLDGELLG-TTQLOACNICIRGKVTGOVPESEQHMWNLEITNLNGTQF 313  
Db 589 NQSHPTTSNAAAPVLDAATGHTNKIQPEDTITRYVQSSQTLDE----- 632  
QY 314 DPTDDVPAPLGVDFEAGEVFGVLQRNGESNPANRAHDV---VATYSDKYTPKGLVQ 370  
Db 633 -----MSVESFLGR-----SGCIHESVLIDVNDYNDQSF TK----- 663

QY 371 IGTWNTDVENOPTKFTPIGLNEVANGHR-FEQWTLPRYSGALTLMNMLAPAVP----- 424  
Db 664 -----WNIN-----LOEMAQIRKRFEMFTYAREFSEIT-----MVPSVAAKDCHI 703  
QY 425 -----LFFGERLLFFERSYVPLKGGFNPDAIDCSVPQEWQHFYQESAPSL----- 469  
Db 704 GHIVMQYMYVPGAPIPTTRDDYAWQSG-----TNASVFVQHGQPPPPRESLPLSLSTASAY 758  
QY 470 -----GDVALVRY---VNPD TG-----RVLFEAKLHK 493  
Db 759 YMFYDGYDGTYSRYGTVTNDMGTLCRSRIVTSEQLHK 797

Search completed: June 1, 2004, 13:47:20  
Job time : 8.52349 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 45.9597 Seconds  
(without alignments)  
3289.030 Million cell updates/sec

Title: US-09-926-799-8  
Perfect score: 2843  
Sequence: 1 MKMASNDAPNSDGAAGLVP.....VNQFYS LAPMTGNGRRRVQ 535

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2843	100.0	535	4 AAB49707	Aab49707 Small rou
2	2269.5	79.8	540	4 AAB49706	Aab49706 Small rou
3	2117.5	74.5	548	4 AAB49705	Aab49705 Small rou
4	2069.5	72.8	548	5 AAU91272	AAU91272 Norwalk v
5	2062.5	72.5	550	4 AAB49709	Aab49709 Small rou
6	2048.5	72.1	542	4 AAB49708	Aab49708 Small rou
7	1919	67.5	539	4 AAB49704	Aab49704 Small rou
8	1656	58.2	541	4 AAB49710	Aab49710 Small rou
9	1200.5	42.2	530	4 AAB49701	Aab49701 Small rou
10	1196.5	42.1	530	2 AAR50972	Aar50972 Norwalk v
11	1196.5	42.1	530	7 ADC72176	Adc72176 Norwalk v
12	1175	41.3	546	4 AAB49702	Aab49702 Small rou
13	1174.5	41.3	545	4 AAB49700	Aab49700 Small rou
14	1151.5	40.5	530	2 AAR57091	Aar57091 Small rou
15	1145	40.3	544	4 AAB49703	Aab49703 Small rou
16	320.5	11.3	579	2 AAW08143	Aaw08143 RHDV caps
17	231	10.2	622	4 AAB47045	Aab47045 Feline Ca
18	230	10.2	623	4 AAB47044	Aab47044 Feline Ca
19	286	10.1	623	4 AAB47043	Aab47043 Feline Ca
20	284	10.0	668	4 AAB67462	Aab67462 Amino aci
21	277.5	9.8	547	4 AAM50108	Aam50108 Feline ca
22	277.5	9.8	669	4 AAB67461	Aab67461 Amino aci
23	277.5	9.8	671	4 AAM50107	Aam50107 Feline ca
24	272.5	9.6	668	2 AARI0686	Aari0686 Feline ca
25	272.5	9.6	668	4 AAE04304	Aae04304 Feline ca

26	194	6.8	40	5 AAU91274	AAU91274 Norwalk v
27	174	6.1	40	5 AAU91273	AAU91273 Norwalk v
28	130	4.6	934	1 AAP20016	Sequence
29	125.5	4.4	1147	5 ABB76724	Foot and
30	125.5	4.4	4643	4 ABB71609	Abb71609 Drosophil
31	122	4.3	980	2 AAW30611	AAW30611 Arabidops
32	122	4.3	980	5 ABB91612	Herbicida
33	122	4.3	985	2 AAW30607	AAW30607 Arabidops
34	121	4.3	761	2 AAR70358	Aar70358 Human NF-
35	117	4.1	338	7 ADC06674	Adc06674 Nervous n
36	114.5	4.0	366	5 ABB05228	Abb05228 Redspotte
37	111	3.9	930	6 ABP96278	Abp96278 Human NF-
38	111	3.9	2206	2 AAR22210	Aar22210 True type
39	110.5	3.9	40	5 AAU91275	AAU91275 Norwalk v
40	110.5	3.9	2468	6 ABU38411	Abu38411 Protein e
41	110.5	3.9	2468	6 ABP59933	Abp59933 Microbial
42	110	3.9	716	2 AAR70357	Aar70357 NF-AT tra
43	110	3.9	716	2 AAW02249	AAW02249 Human tra
44	110	3.9	716	3 AAY96565	Aay96565 Human nuc
45	110	3.9	716	3 AAY96535	Aay96535 Human nuc

ALIGNMENTS

RESULT 1  
AAB49707  
ID AAB49707 standard; protein; 535 AA.

AC AAB49707;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 8.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.

XX WO200079280-A1.

XX 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.

XX 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX WPI; 2001-080848/09.

DR N-PSDB; AAF29148.

XX Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

XX Claim 1; Page 57-59; 84pp; Japanese.

PS This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

SQ Sequence 535 AA;

Query Match 100.0%; Score 2843; DB 4; Length 535;  
Best Local Similarity 100.0%; Pred. No. 2.2e-252;

Matches	535;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;	
Y	1	MMKASNDAA	PSNDGAAGLVPEANN	ETMALEPVAGASIA	APLTGQNNIIDPWIRLNF	VOAP	60			
b	1	MMKASNDAA	PSNDGAAGLVPEANN	ETMALEPVAGASIA	APLTGQNNIIDPWIRLNF	VOAP	60			
Y	61	NGEFTVSP	RNSPGEVLLNLELGP	ELNPYLALHLSRM	YNGVAGVEVQVLLAGNAFTAG	KLIV	120			
b	61	NGEFTVSP	RNSPGEVLLNLELGP	ELNPYLALHLSRM	YNGVAGVEVQVLLAGNAFTAG	KLIV	120			
Y	121	FAAVPPH	PLENISPGQITMFP	HVIDVRLTSPVLLPLPDVVRNNF	FHYNQNEPRMLVA	180				
b	121	FAAVPPH	PLENISPGQITMFP	HVIDVRLTSPVLLPLPDVVRNNF	FHYNQNEPRMLVA	180				
Y	181	MLYTP	LRNSGDDVFTV	SCRVLTRSPDFDNVLVPP	TLESKTPFTLPILITIGELTNS	240				
b	181	MLYTP	LRNSGDDVFTV	SCRVLTRSPDFDNVLVPP	TLESKTPFTLPILITIGELTNS	240				
Y	241	RFPVP	IDELYTS	PNESLVVQPNQGR	CALDGELOQTQLLPTAICS	FRGRINOKVGENHV	300			
b	241	RFPVP	IDELYTS	PNESLVVQPNQGR	CALDGELOQTQLLPTAICS	FRGRINOKVGENHV	300			
Y	301	WNMQVT	NINPTPTD	PTGDVPAPLGTDFSGKLF	GVLSQRDHONACKSHDAVI	ATNSAKFT	360			
b	301	WNMQVT	NINPTPTD	PTGDVPAPLGTDFSGKLF	GVLSQRDHONACKSHDAVI	ATNSAKFT	360			
Y	361	PKLGA	IOIGTWEEDDVH	INOPTKFTVPGLF	ENEGFNQWTLPNYSGALT	LNMG	LAPVAPT	420		
b	361	PKLGA	IOIGTWEEDDVH	INOPTKFTVPGLF	ENEGFNQWTLPNYSGALT	LNMG	LAPVAPT	420		
Y	421	FPGEQIL	FFRSHIPLK	GGVADPVIDCLL	POEWI	QHLYQESAPSQSD	VALIRFTN	PD	TGVR	480
b	421	FPGEQIL	FFRSHIPLK	GGVADPVIDCLL	POEWI	QHLYQESAPSQSD	VALIRFTN	PD	TGVR	480
Y	481	LFEAK	LHRS	GYITVANTGSR	PIVVPANGYFR	EDTWNQFYS	LAPMG	TGNGRRVQ	535	
b	481	LFEAK	LHRS	GYITVANTGSR	PIVVPANGYFR	EDTWNQFYS	LAPMG	TGNGRRVQ	535	
RESULT 2										
AAB49706										
ID	AAB49706 standard; protein; 540 AA.									
XX	AAB49706;									
DT	04-APR-2001 (first entry)									
XX	Small round structured virus protein SEQ ID 7.									
XX	Small round structured virus; SRSV; food poisoning.									
XX	Small round structured virus.									
OS	WO200079280-A1.									
PN	28-DEC-2000.									
PD	22-JUN-2000; 2000WO-JP004095.									
PF	22-JUN-1999; 99JP-00175928.									
PR	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.									
XX	(DENK-) DENKA SEIKEN KK.									
PA	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;									
XX	WPI: 2001-080848/09.									
DR	N-PSDB; AAF29147.									
XX	Kit for the detection and typing of small round-structured virus (SRSV)									
PT	strains for investigation of food poisoning outbreaks, contains									
PT	antibodies.									
XX										

Claim 1; Page 54-57; 84pp; Japanese.

PS This invention relates to a kit for the detection and typing of small  
XX round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX

SQ Sequence 540 AA;

Query Match 79.8%; Score 2269.5; DB 4; Length 540;

Best Local Similarity 76.3%; Pred. No. 1.4e-199;

Matches 413; Conservative 65; Mismatches 56; Indels 7; Gaps 3;

QY	1	MMKASNDAA	PSNDGAAGLVPEANN	ETMALEPVAGASIA	APLTGQNNII	DPWIRLNF	VOAP	60			
Db											
QY	1	MMKASNDAA	PSNDGAAGLVPEANN	ETMALEPVAGASIA	APLTGQNNII	DPWIRLNF	VOAP	60			
Db											
QY	61	NGEFTVSP	RNSPGVLLNLELGP	ELNPYLALHLSRM	YNGVAGVEVQV	LLAGNAFTAG	KLIV	120			
Db											
QY	61	NGEFTVSP	RNSPGVLLNLELGP	ELNPYLALHLSRM	YNGVAGVEVQV	LLAGNAFTAG	KLIV	120			
Db											
QY	121	FAAVPPH	PLENISPGQITMFP	HVIDVRLTSPVLL	PLPDVVRNNF	FHYNQNEPR	MLVA	180			
Db											
QY	121	FAAVPPH	PLENISPGQITMFP	HVIDVRLTSPVLL	PLPDVVRNNF	FHYNQNEPR	MLVA	180			
Db											
QY	181	MLYTP	LRNSGDDVFTV	SCRVLTRSPDP	DFNVLVPP	TLESKTPFTL	PILITIGELTNS	240			
Db											
QY	181	MLYTP	LRNSGDDVFTV	SCRVLTRSPDP	DFNVLVPP	TLESKTPFTL	PILITIGELTNS	240			
Db											
QY	241	RFPVP	IDELYTS	PNESLVVQ	PQNGRCALD	GELQGT	TQLLPTAICS	FRGRINOKV	GENHV 300		
Db											
QY	241	RFPVP	IDELYTS	PNESLVVQ	PQNGRCALD	GELQGT	TQLLPTAICS	FRGRINOKV	GENHV 300		
Db											
QY	301	WNMQV	TNINPT	PTDGDVP	APLGTDF	SGKLF	GVLSQR	---DHDNACH	SHDAVIATNSA 357		
Db											
QY	301	WNMQV	TNINPT	PTDGDVP	APLGTDF	SGKLF	GVLSQR	---DHDNACH	SHDAVIATNSA 357		
Db											
QY	358	KFTPK	GAIOIGT	WEEDDVH	INOPTKFT	VPGLF	---NEGFNQ	WTLPNYSGALTL	NMG	LA 414	
Db											
QY	358	KFTPK	GAIOIGT	WEEDDVH	INOPTKFT	VPGLF	---NEGFNQ	WTLPNYSGALTL	NMG	LA 414	
Db											
QY	415	PPVAP	TFE	QILFRSHI	PLKGGVADP	VIDCLLP	POEWI	QHLYES	APSQSD	VALIRFTN 474	
Db											
QY	415	PPVAP	TFE	QILFRSHI	PLKGGVADP	VIDCLLP	POEWI	QHLYES	APSQSD	VALIRFTN 474	
Db											
QY	475	PDTGR	VLF	PEAKLHRS	GYITVANT	GSPIVVP	PANGYFR	EDTWNQ	FYS	LAPMG	TGNGRRRV 534
Db											
QY	475	PDTGR	VLF	PEAKLHRS	GYITVANT	GSPIVVP	PANGYFR	EDTWNQ	FYS	LAPMG	TGNGRRRV 534
Db											
QY	535	Q	535								
Db											
QY	540	Q	540								
Db											

RESULT 3

AAB49705  
ID AAB49705 standard; protein; 548 AA.

XX AAB49705;

XX 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 6.

XX Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.

XX WO200079280-A1.

XX



```
PD 28-DEC-2000.
XX
XX PF 22-JUN-2000; 2000WO-JF004095.
XX PR 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX PA (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI: 2001-080848/09.
XX DR N-PSDB; AAP29146.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX
XX Claim 1; Page 52-54; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX CC round structured virus (SRSV) strains. The kit contains antibodies
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAP20141 -
XX CC AAP20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 548 AA;
SQ
Query Match 74.5%; Score 2117.5; DB 4; Length 548;
Best Local Similarity 70.5%; Pred. No. 1.4e-185;
Matches 387; Conservative 74; Mismatches 73; Indels 15; Gaps 4;
QY 1 MKMASNDRAAPSDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
DB 1 MKMASNDRAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNNIIDPWIRLNFVQAP 60
QY 61 NGFEVTSRNSPGEVLLNLELGPENLPYLAHLARMYNGYAGVEVQVLLAGNAFTAGKLV 120
DB 61 GGEFTVSPNSPGEVLLNLELGPENLPYLAHLARMYNGYAGVEVQVLLAGNAFTAGKII 120
QY 121 FAAPVPHPLENISPGQITMFPVHVIDVRLPEVLLPDPVRNPFHYNQNEPRMLVA 180
DB 121 FAAPVPHPLENISPGQITMFPVHVIDVRLPEVLLPDPVRNPFHYNQNEPRMLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDFNVLVPPTLESKTPFTLPILTIGELTNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDFNVLVPPTLESKTPFTLPILTIGELTNS 240
QY 241 RFPVPIDELTSPNESLVVQONGRCALDGELOQTTLPTAICSPGRINQKVSGE--- 297
DB 241 RFPVPIDELTSPNESLVVQONGRCALDGELOQTTLPTAICSPGRINQKVSGE--- 297
QY 301 ADTATPRLFNYYWHVQDLNLTGTPYDPAEDIPGLTDPFRGKVGVSQRNLDSTTRAH 360
DB 301 ADTATPRLFNYYWHVQDLNLTGTPYDPAEDIPGLTDPFRGKVGVSQRNLDSTTRAH 360
QY 349 DAVIATNSAKFTPKLGAIGITWEEDVHINQKTPVGL- FENEG- FNQWTLNYSGA 406
DB 349 DAVIATNSAKFTPKLGAIGITWEEDVHINQKTPVGL- FENEG- FNQWTLNYSGA 406
QY 361 EAKVDITAGRTFTPKLGSLSIST- DSDDFDQONQKTPFTVGVIGDNEAEFQWLSLDYSGQ 419
DB 361 EAKVDITAGRTFTPKLGSLSIST- DSDDFDQONQKTPFTVGVIGDNEAEFQWLSLDYSGQ 419
QY 407 LTLNMLGAPVAPPTFPGEOILFPRSHI PLKGGVADPVIDCLLPQEWIQLYQESAPSQSD 466
DB 407 LTLNMLGAPVAPPTFPGEOILFPRSHI PLKGGVADPVIDCLLPQEWIQLYQESAPSQSD 466
QY 420 FTHNMMLAFAVAPNPFGEQQLLFRSOLPSSGGRSNGVLDCLVPQEWQHFYQESAPATQ 479
DB 420 FTHNMMLAFAVAPNPFGEQQLLFRSOLPSSGGRSNGVLDCLVPQEWQHFYQESAPATQ 479
QY 467 VALIRFTNPDGTGRVLFPEAKLHRSGYITVANTGSRPIWVPANGYFRPDTWVNOFYSLAPMG 526
DB 467 VALIRFTNPDGTGRVLFPEAKLHRSGYITVANTGSRPIWVPANGYFRPDTWVNOFYSLAPMG 526
QY 480 VALVRYVNPDTGKVLFEAKLHLKLGFTIANNGDSPIITVPPNGYFRFESWVNFYTLAPMG 539
DB 480 VALVRYVNPDTGKVLFEAKLHLKLGFTIANNGDSPIITVPPNGYFRFESWVNFYTLAPMG 539
QY 527 TNGRRRRVQ 535
DB 527 TNGRRRRVQ 535
QY 540 TNGRRRRQ 548
DB 540 TNGRRRRQ 548
RESULT 4
AAU91272
ID AAU91272 standard; protein; 548 AA.
XX
XX AC AAU91272;
XX
XX DT 18-JUN-2002 (first entry)
XX
XX Norwalk virus associated polynucleotide #1.
XX DE Norwalk virus; monoclonal antibody; geno group I; geno group II;
XX KW immunological detection; food; viral infection.
XX KW Norwalk virus.
XX OS Norwalk virus.
XX
XX JP2002020399-A.
XX
XX PD 23-JAN-2002.
XX
XX PF 10-JUL-2000; 2000JP-00208151.
XX
XX PR 10-JUL-2000; 2000JP-00208151.
XX
XX PA (OSAP) OSAKA PREFECTURE.
XX PA (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.
XX PA (IATR) IATRON LAB INC.
XX
XX WPI: 2002-287412/33.
XX
XX A monoclonal antibody useful in the immunological detection and diagnosis
XX PT of Norwalk virus infection.
XX
XX Disclosure; Page 12-13; 24pp; Japanese.
XX
XX The invention describes a monoclonal antibody recognising Norwalk virus,
XX CC a capsid protein of Norwalk virus, or a common antigen epitope on the
XX CC capsid protein molecule of geno group I and geno group II. The antibody
XX CC is useful for immunological detection and quantitative analysis of
XX CC Norwalk virus in foods and the serum of infected patients. This sequence
XX CC represents a Norwalk virus associated protein described in the invention
XX
XX Sequence 548 AA;
SQ
Query Match 72.8%; Score 2069.5; DB 5; Length 548;
Best Local Similarity 69.7%; Pred. No. 3.7e-181;
Matches 380; Conservative 73; Mismatches 77; Indels 15; Gaps 4;
QY 1 MKMASNDRAAPSDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
DB 1 MKMASNDRAAPSDGAGLVPEINNEAMALDPVAGAAIAAPLTGQNNIIDPWIRLNFVQAP 60
QY 61 NGFEVTSRNSPGEVLLNLELGPENLPYLAHLARMYNGYAGVEVQVLLAGNAFTAGKLV 120
DB 61 GGEFTVSPNSPGEVLLNLELGPENLPYLAHLARMYNGYAGVEVQVLLAGNAFTAGKII 120
QY 121 FAAPVPHPLENISPGQITMFPVHVIDVRLPEVLLPDPVRNPFHYNQNEPRMLVA 180
DB 121 FAAPVPHPLENISPGQITMFPVHVIDVRLPEVLLPDPVRNPFHYNQNEPRMLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDFNVLVPPTLESKTPFTLPILTIGELTNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDFNVLVPPTLESKTPFTLPILTIGELTNS 240
QY 241 RFPVPIDELTSPNESLVVQONGRCALDGELOQTTLPTAICSPGRINQKVSGE--- 297
DB 241 RFPVPIDELTSPNESLVVQONGRCALDGELOQTTLPTAICSPGRINQKVSGE--- 297
QY 241 RFPVPIDELTSPNESLVVQONGRCALDGELOQTTLPTAICSPGRINQKVSGE--- 297
DB 241 RFPVPIDELTSPNESLVVQONGRCALDGELOQTTLPTAICSPGRINQKVSGE--- 297
QY 298 -----NHVNMVQNTINGTPTDPTGDPAPLGTDPFGKLVSGKLVSGRDNACRSH 348
DB 298 -----NHVNMVQNTINGTPTDPTGDPAPLGTDPFGKLVSGKLVSGRDNACRSH 348
QY 301 ADIATPRLFNYYWHVQDLNLTGTPYDPAEDIPGLTDPFRGKVGVSQRNLDSTTRAH 360
DB 301 ADIATPRLFNYYWHVQDLNLTGTPYDPAEDIPGLTDPFRGKVGVSQRNLDSTTRAH 360
```



CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX  
 XX Sequence 542 AA;

Query Match 72.1%; Score 2048.5; DB 4; Length 542;  
 Best Local Similarity 71.7%; Pred. No. 3.1e-179;  
 Matches 392; Conservative 64; Mismatches 74; Indels 17; Gaps 7;  
 QY 1 MKMASNDAPSDNDGAGLVPEANNTMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60  
 DB 1 MKMASNDAPSDNDGAGLVPEANNTMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60  
 QY 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVNGYAGVEVQVLLAGNAFTAGKLV 120  
 DB 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVNGYAGVEVQVLLAGNAFTAGKLV 120  
 QY 121 FAAPVPHPLENISPGQITMFPVHVIDVRLPVLPLPDVNNFFHYNQNEPRMLVA 180  
 DB 121 FAAPVPHPLENISPGQITMFPVHVIDVRLPVLPLPDVNNFFHYNQNEPRMLVA 180  
 QY 181 MLYTFLRNSGDDVFTVSCRVLTRPSDPDFNVLPPPTLESKTKPFTLPILTIGELTNS 240  
 DB 181 MLYTFLRNSGDDVFTVSCRVLTRPSDPDFNVLPPPTLESKTKPFTLPILTIGELTNS 240  
 QY 241 RPPVPIDELYTSPNESLVVQPNQGRCALDGELOQTQLLPTAICSFGRINQKVSNGHV 300  
 DB 241 RPPVPIDELYTSPNESLVVQPNQGRCALDGELOQTQLLPTAICSFGRINQKVSNGHV 300  
 QY 301 WN---MQVTNINGTPTDPTGDPVAPLGTDPFSGKLVGLVLSQRDHDN-----ACRSHDAVI 352  
 DB 298 WDNLLQLTYPNGASVDPDTPVAPLGTDPFSGKLVGLVLTQ-DNVNVSNGEAKNAKGI 356  
 QY 353 ATNSAKFTPKLGAIGTWEEDVHINQPTKTPVGLF--ENEGFNQWTLNYSALTN 410  
 DB 357 STTSKFTPKIGISGLHSITE-HVHFNQOSRETFVGVAVDENTPPQWVLPYAGSLAN 415  
 QY 411 MGLAPVAPTPGEOILFFRSHIPLKGGV--ADPVIDCLLPQEWIOLYQESAPQSQDVA 468  
 DB 416 TMLAPAVAPTPGEOILFFRSHIPLKGGV--ADPVIDCLLPQEWIOLYQESAPQSQDVA 475  
 QY 469 LIRFTNPDGTGRVFEAKLHRSYITVANTGSRPIVVPVANGYFRFDTWVNFQYSLAPMG 528  
 DB 476 LIRVYNDPTGRVFEAKLHRSYITVANTGSRPIVVPVANGYFRFDTWVNFQYSLAPMG 535  
 QY 529 NGRRRVQ 535  
 DB 536 NGRRRRIQ 542

RESULT 7  
 AAB49704  
 ID AAB49704 standard; protein; 539 AA.

XX AC AAB49704;

XX DT 04-APR-2001 (first entry)

XX DE Small round structured virus protein SEQ ID 5.

XX KW Small round structured virus; SRSV; food poisoning.

XX OS Small round structured virus.

XX PN WO200079280-A1.

XX PD 28-DEC-2000.

XX XX 22-JUN-2000; 2000WO-JP004095.

XX

PR 22-JUN-1999; 99JP-00175928.

XX PA (NTNA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 XX PA (DENK-) DENKA SEIKEN KK.

XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

XX DR WPI; 2001-080848/09.

XX DR N-PSDB; AAF29145.

XX PT Kit for the detection and typing of small round-structured virus (SRSV)  
 XX PT strains for investigation of food poisoning outbreaks, contains  
 XX PT antibodies.

XX PS Claim 1; Page 50-52; 84pp; Japanese.

XX CC This invention relates to a kit for the detection and typing of small  
 XX CC round structured virus (SRSV) strains. The kit contains antibodies  
 XX CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 XX CC used for detecting and typing strains of SRSV in order to prevent the  
 XX CC spread of infection and to examine the epidemiology of outbreaks

XX SQ Sequence 539 AA;

Query Match 67.5%; Score 1919; DB 4; Length 539;  
 Best Local Similarity 65.4%; Pred. No. 2.6e-167;

Matches 353; Conservative 80; Mismatches 97; Indels 10; Gaps 5;

QY 1 MKMASNDAPSDNDGAGLVPEANNTMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60  
 DB 1 MKMASNDAPSDNDGAGLVPEANNTMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60  
 QY 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVNGYAGVEVQVLLAGNAFTAGKLV 120  
 DB 61 GGEFTVSPRNPAGGILLWSAPLCPDLNPLYLHARMYNGYAGVEVQVLLAGNAFTAGKII 120  
 QY 121 FAAPVPHPLENISPGQITMFPVHVIDVRLPVLPLPDVNNFFHYNQNEPRMLVA 180  
 DB 121 FAAPVPHPLENISPGQITMFPVHVIDVRLPVLPLPDVNNFFHYNQNEPRMLVA 180  
 QY 181 MLYTFLRNSGDDVFTVSCRVLTRPSDPDFNVLPPPTLESKTKPFTLPILTIGELTNS 240  
 DB 181 MLYTFLRANAGDDVFTVSCRVLTRPSDPDFNVLPPPTLESKTKPFTLPILTIGELTNS 240  
 QY 241 RPPVPIDELYTSPNESLVVQPNQGRCALDGELOQTQLLPTAICSFGRINQKVSNGHV 300  
 DB 241 RPPVPIDELYTSPNESLVVQPNQGRCALDGELOQTQLLPTAICSFGRINQKVSNGHV 300  
 QY 301 WNMQVTNINGTPTDPTGDPVAPLGTDPFSGKLVGLVLSQRD-HDNACRSHDAVIATNSAKF 359  
 DB 299 YTMNLASQNSNYDPTBEIPAPLGTDPFVGIQGLMTQTTREDGSTRAHKATVSTGVHP 358  
 QY 360 TPKLGAIGTWEEDVHINQPTKTPVGLFENGFN-----QWTLNYSALTNLGL 413  
 DB 359 TPKLGSVQYTTDNTNDFQTGQTKFTFPVGVITQ-DGNHQNQEPQWVLPYNSGRTGHVHL 417  
 QY 414 APPVAPTPGEOILFFRSHIPLKGGVADPVIDCLLPQEWIOLYQESAPQSQDVALIRFT 473  
 DB 418 APAVAPTPGEOILFFRSTMPGCSGYPMNNDCLLPQEWVQHFQOEAPAPSDVALLRFV 477  
 QY 474 NPDTRVLFEAKLHRSYITVANTGSRPIVVPVANGYFRFDTWVNFQYSLAPMGNGRRR 533  
 DB 478 NPDTRVLFECKLHKSIVTVHTGPHDLVTPPNGYFRFDSWVNFQYTLAPMGNGAGRRR 537

RESULT 8

AAB49710

ID AAB49710 standard; protein; 541 AA.

XX AC AAB49710;

XX

```

DT 04-APR-2001 (first entry)
XX Small round structured virus protein SEQ ID 11.
DE Small round structured virus; SRSV; food poisoning.
KW Small round structured virus; SRSV; food poisoning.
XX Small round structured virus.
OS Small round structured virus.
XX WO200079280-A1.
XX 28-DEC-2000.
XX 22-JUN-2000; 2000WO-JP004095.
XX 22-JUN-1999; 99JP-00175928.
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI; 2001-080848/09.
XX N-PSDB; AAF29151.
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX Claim 1; Page 64-66; 84pp; Japanese.
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 541 AA;
XX
XX Query Match 58.2%; Score 1656; DB 4; Length 541;
XX Best Local Similarity 59.2%; Pred. No. 4.2e-143;
XX Matches 330; Conservative 67; Mismatches 122; Indels 38; Gaps 7;
XX
QY 1 MKMASNDAAAPSDGAGLVPEANNETWALBPVAGASTAAPLTGQNNIIDPWIRNFVQAP 60
DB 1 MKMASNDAAAPSDGAGLVPEANNETWALBPVAGASTAAPLTGQNNIIDPWIRNFVQAP 60
QY 61 NGFETVSPRNSPGEVLLNLELGPENLPYLAHLSRMYNGYAGGVEVQVLLAGNAFTAGKLV 120
DB 61 AGEFTVSPRNSPGEVLLNLELGPENLPYLAHLARMYNGHAGGVEVQVLLAGNAFTAGKII 120
QY 121 FAAVPPHPPLENISPQGITMFPFVHIIIVRTLEPVLPLPDVRRNFFHYNQNEPRMLVA 180
DB 121 FAAPPPGPFYENLSPSQITMCPHFVHIIIVRQLEPVLPLPDVRRNFFHYNQNEPRMLVA 180
QY 181 MLYTLRNGSGDDVFTVSCRVLPDPFNVLPVETLESKTPFTLITLIGELTNS 240
DB 181 MLYTLRANSGDDVFTVSCRVLPDPFNVLPVETLESKTPFTLITLIGELTNS 240
QY 241 RFPVPIDELTYSPNESLVVQNGRCALDGLQGTQLLPTAICSFRRINQKVSGB-N 298
DB 241 RFPVPDVVYIARNENQVQNGRVTLDGELLGTTPLAVNICFKFGEVIK-NGDVRS 299
QY 299 HYWNQVNTNINGTTPDPTGDVPAPLGTDFSGKLFVLSQR----DHDNACHSHDAVTATN 355
DB 300 YRMDMEITNDCTPIDPTDPTGPTGSPDFQGLEFGVASQRNKNQNEQNPATRAHEAINTG 359
QY 356 SAKFTPKLGAIGICITWEDDVHINQTKPTPGLFENEGNCQWTLPNYSGALTL----- 409
DB 360 GDHLCPQISSSEIYITSPNLCRTNPQPLQPSGL-----RGTLIRSDNGH 405
QY 410 --NMGLAPPVAPTFPGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSQSD- 466

```

RESULT 9

```

AAB49701
ID AAB49701 standard; protein; 530 AA.
XX
AC AAB49701;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 2.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
FN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 99JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX
PA (DENK-) DENKA SEIKEN KK.
XX
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
WPI; 2001-080848/09.
XX
N-PSDB; AAF29142.
XX
Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX
Claim 1; Page 42-45; 84pp; Japanese.
XX
This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 530 AA;
XX
XX Query Match 42.2%; Score 1200.5; DB 4; Length 530;
XX Best Local Similarity 46.4%; Pred. No. 3.5e-101;
XX Matches 258; Conservative 76; Mismatches 165; Indels 57; Gaps 12;
XX
QY 1 MKMASNDAAAPSDGAG---LVPEAN-NEFWALEPVAGASTAAPLTGQNNIIDPWIRLNF 56
DB 1 MMMAKSDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGQNPIDPWIRLNF 60
QY 57 VQAPNGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMYNGYAGGVEVQVLLAGNAFTA 116
DB 61 VQAPQGEFTISPNNITFGGVLFDLGLGPHLPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
QY 117 GKLVFAAVPPHPPLENISPQGITMFPFVHIIIVRTLEPVLPLPDVRRNFFHYNQNEPRM 176
DB 121 GKLIIVSCIPPGFGSHNLITIAQTALFPHVITADVRLTDPIEVPLEDVRRNVLFFHNDNRNQQT 180

```

Qy		177	RLVAMLYTPLRSGDGDVFVTVCSEVLTRPSDEPDFNVLVPPTLESKTKPFTLPILTIGE	236
Dd		181	RLVCMVLTPLRTGGTGSFVVAGRVMTCPSPDENFLFVLVPPTVEQKRPTFTLPLNLSS	240
Qy		237	LTSRFPVPIDELYSNPESLVQPONGRCALDGELOQTQLLPTAI CSFRGRINQKVSG	296
Dd		241	LSNRAPLLPIGSMGISPNVQSVPQNGRCRTL DGLRVGTTPVSLSHVAKIRGTSNGTV--	298
Qy		297	ENHVNNMQVTNINGTFPDPTGDVPAPLCTPDFSG-----KLFGVLSQRDHDAACRSKD	349
Dd		299	-----TNLTEDGTPPHPP-EGPAPIGFDPDLGGCDWHINTQFHSSQTQYD-----	344
Qy		350	AVIATNSAKFTPKLGAIQ-----IG--TWEEDDVHIHQPTFTPVGLFENEGFNOW	398
Dd		345	--VDTTPTDTFFPHLGSIQAINGISGNYLGVLSWSPSSH-----PSG-----SQVDLM	390
Qy		399	TLPNYSGALTLMNGLAPPVAPTFFQEOLFFRRSHIP LKGGVADPIDCLLPQEWIOHLYQ	458
Dd		391	KIPNYGSSITEATHLAPSVPYPGFEVLVFFMMSKIPPGGAYSLP---CLLPQEYISHLAS	447
Qy		459	ESAPSQSDVALIRNTNPDGTGRVLFEAKLHRSGYITVAN TG--SRPIVVDPANGYFPFDTWV	516
Dd		448	EQAPTUGEALLHYVDPDGTGRLTGEFKAYPDGFLT CPVGASSGQQLPINGVFPVFSWV	507
Qy		517	NQFYSLAPMGTCNGRR	532
Dd		508	SRYQLKPVGVTASSAR	523

**RESULT 10**

AAAR50972	
ID	AAAR50972 standard; protein; 530 AA.
XX	
XX	AAAR50972;
XX	
XX	16-OCT-2003 (revised)
DT	25-MAR-2003 (revised)
DT	05-OCT-1994 (first entry)
XX	
XX	Norwalk virus strain 8FIIa protein (encoded by ORF2).
DE	
XX	
KW	Norwalk virus; pathogen; acute gastroenteritis; food poisoning;
KW	seafood contamination; diagnostic assay; calcivirus; small round virus.
XX	
OS	Norwalk virus; (strain 8FIIa).
XX	
PN	WO9405700-A2.
XX	
PD	17-MAR-1994.
XX	
XX	07-SEP-1993; 93WO-US008447.
PF	
XX	
PR	07-SEP-1992; 92US-00941365.
XX	
PA	(BAYU ) BAYLOR COLLEGE MEDICINE.
XX	
PI	Matson DO, Estes MK, Jiang X, Graham DY;
XX	
DR	WPI; 1994-101125/12.
DR	N-PSDB; AAQ56826.
XX	
XX	DNA from Norwalk and related viruses - used for preparing prods. for use
PT	in diagnostic assays, detection and vaccines for Norwalk and related
PT	viruses.
XX	
PS	Claim 14; Page 68-70; 156pp; English.
XX	
CC	The Norwalk virus was isolated from stool samples from adult volunteers
CC	infected with safety tested Norwalk virus strain 8FIIa. the coding
CC	sequence is useful for the design of probes for use in diagnostic assays
CC	for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct
CC	PN field.) (Updated on 16-OCT-2003 to standardise OS field)



```

Db 352 SVQTNVQGVFVPHLGSIQDEVENHHTGDYIGIE-----WISQPS--TPPG-----TDINLW 401
Qy 399 TLPNYSGALTLMNGLAPPVAPTFPGGEILFFRSHIPLKGGVADP-VIDCLLPQEWIOHLY 457
Db 402 EIPDYGSSLSQAAANLAPPVFPFGFGEALVYFVSAPGPNRSAPNDVPCLLPQEVITHFV 461
Qy 458 QESAPSQSDVALIRTNPDTRGVLPFAKLRHSGYITVA--NTGSRPIVVPANGYFRFTDW 515
Db 462 SEQAPTMGDAALLHYVDPDTRNLGFEKLYPGCYLTCVPGNGVAGPQQLPLNGVLFVSW 521
Qy 516 VNQFYSLAPMGTCN-----GRRV 534
Db 522 VSRFYQLKPVGTASTARSRLGVRI 546

RESULT 13
AAB49700
ID AAB49700 standard; protein; 545 AA.
XX
AC AAB49700;
XX
DT 04-APR-2001 (first entry)
XX
DE Small round structured virus protein SEQ ID 1.
XX
KW Small round structured virus; SRSV; food poisoning.
XX
OS Small round structured virus.
XX
PN WO200079280-A1.
XX
PD 28-DEC-2000.
XX
PF 22-JUN-2000; 2000WO-JP004095.
XX
PR 22-JUN-1999; 95JP-00175928.
XX
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
WPI; 2001-080848/09.
DR N-PSDB; AAF29141.
XX
PT Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
PS Claim 1; Page 40-42; 84pp; Japanese.
XX
CC This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
SQ Sequence 545 AA;

Query Match 41.3%; Score 1174.5; DB 4; Length 545;
Best Local Similarity 45.0%; Pred. No. 9e-99;
Matches 251; Conservative 87; Mismatches 181; Indels 39; Gaps 13;

Qy 1 MKWASNDAPSNDGAG---LYPEANN-ETMALEPVAGASIAAPLTGQNNIIDPWIRLNF 56
Db 1 MMWASKADPTNMDTSGAQQLVPEANTAEPISEMPVAGATAAATAGQVNMIDPWIMNY 60
Qy 57 VQAPNGEFTVSPRSPGVELLMLELGPENLPVLAHLSRMVNGYAGGVEQVLLAGNAFTA 116
Db 61 VQAPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLAQWYNGVGNMKYKVLLAGNAFTA 120

```

```

Qy 117 GKLFAAAPPHPHLENISPGQITMPHPHVIDVRLTEPVLPLDVRNNPFFHNOQNEPRM 176
Db 121 KIIISCIIPPGFAAQNISIAQATMPHPHVIADVRLTEPVLPLDVRNNPFFHNOQNEPRM 179
Qy 177 RLVAMLYTPPLASNG--SGDDVFTVSCRVLTRPSDPDFENYLVPPPTLESKTKPFTLPILT 234
Db 180 RLVCMLYTPPLASNGSGSGTDFPVIAGRVLTCPSPDFSLFLVPPNVEOKTKPSPVNDPL 239
Qy 235 GELTNSRFPVTDILYTSFNESLVVQPQNGRCALDGLQGTTLQLLPTAICSPGRINQKV 294
Db 240 NTLNSRVPSLTKSMVSRDHQWVQFQNGRVTLQGLQGTTPTSASQLCKIRGSVFHAN 299
Qy 295 SGENHVMNQVTNINGTPTDGVDPAPLGTDPFGKLFGLVLSQSD-HDNAC-----RSH 348
Db 300 GNGY----NLTELDGSPYH-AFESAPAPIGFPD-----LGECDWHMEASPTQFNTG 346
Qy 349 DAVIATN---SAKFTPKLGAIGTWEEEDDVHINQTKF---TPVGLPENEGFNQWTLN 402
Db 347 DVIKQINVKQBSAFAPHLGTIQADGLSDVSVNTNMIAGLWVSPVSDGHRGVDVPIPR 406
Qy 403 YSGALTLMNGLAPPVAPTFPGGEILFFRSHIPLKGGVADPVIDCLLPQEWIOHLYQESAP 462
Db 407 YGSTLTEAAQLAPPIYPFGFGEAIVFFMSDFPIAHGTNGLSVPCCTIPQEFVTHFVNEQAP 466
Qy 463 SQSDVALIRTNPDTRGVLPFAKLRHSGYITVA--NTGSRPIVVPANGYFRFTDWVNOQY 520
Db 467 TRGEAALLHYLDPDTRNLGFEKLYPGCYLTCVPGNGVAGPQQLPLNGVLFVSWVREF 526
Qy 521 SLAPMGTCN-----GRRR 533
Db 527 QLKPVGTAGPACRLGIRR 544

RESULT 14
AAR57091
ID AAR57091 standard; protein; 530 AA.
XX
AC AAR57091;
XX
DT 27-AUG-2003 (revised)
DT 25-MAR-2003 (revised)
DT 05-OCT-1994 (first entry)
XX
DE Small round virus SRSV/KY/89 capsid protein.
XX
KW pathogen; acute gastroenteritis; food poisoning; seafood contamination;
KW diagnostic assay; human calicivirus; small round virus; SRSV; KY89;
KW Norwalk virus; capsid protein.
XX
OS Small round structured virus.
XX
PN WO9405700-A2.
XX
PD 17-MAR-1994.
XX
PF 07-SEP-1993; 93WO-US008447.
XX
PR 07-SEP-1992; 92US-00941365.
XX
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI Matson DO, Estes MK, Jiang X, Graham DY;
XX
WPI; 1994-101125/12.
DR N-PSDB; AAQ56832.
XX
PT DNA from Norwalk and related viruses - used for preparing prods. for use
PT in diagnostic assays, detection and vaccines for Norwalk and related
PT viruses.
XX
XX Example 7; Fig 13a; 156pp; English.
XX

```

CC The known sequence for Norwalk virus was used to obtain the sequence of  
 CC other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool  
 CC from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide  
 CC cDNA sequence includes part of the polymerase region and the capsid  
 CC region of the genome; the deduced amino acid sequences are AAR57092 and  
 CC AAR57091, respectively. Expression of fragments and derivs. of Norwalk-  
 CC related viruses permits development of diagnostic assays to detect  
 CC antibodies, antigens, viral genetic material or antivirals. (Updated on  
 CC 25-MAR-2003 to correct FN field.) (Updated on 27-AUG-2003 to correct OS  
 CC field.)  
 XX  
 XX

XX Sequence 530 AA;

Query Match 40.5%; Score 1151.5; DB 2; Length 530;  
 Best Local Similarity 45.4%; Pred. No. 1.1e-96;  
 Matches 253; Conservative 79; Mismatches 166; Indels 59; Gaps 14;

QY 1 MKMASNDAAAPSDGAG---LVPEAN-NETMALEPVAGASIAAPLTGQNNIIDPWIRLNF 56  
 DB 1 MMASKDATSSVDGASASQVLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINF 60  
 QY 57 VQAPNGEFTVSPRNSPCEVLLNLELGPENLPYLAHLRMYNGYAGGVEQVLLAGNAFTA 116  
 DB 61 VQAPNGEFTISPNTTSGDVLFDLSLGFHLNPFLLHLSQYNGWGMVRVIMLAGNAFTA 120  
 QY 117 GKLVFAAVPPHFFLENIISPGQITMFPVHVIDVRLTEPVLPLPDVNRNFFHYNQNEPRM 176  
 DB 121 GKLIIVCIPFGSGQQLTIAQATLFPVHVIDVRLTEPVLPLPDVNRNFFHYNQNEPRM 180  
 QY 177 RLVMAMLYTPURSGSGDDVFTVSCRVLTRPSDFDNLYVPPTLESKTKPFTLPILTIGE 236  
 DB 181 RLVCMLYTPURSGSGDDVFTVSCRVLTRPSDFDNLYVPPTLESKTKPFTLPILTIGE 240  
 QY 237 LTNRSRFPVPIDELYTSPNESLVVQNGRCALDGLQGTTLQTLTAICFRGRINOKV 296  
 DB 241 LNSNRAPLPISGMISFDNVQSVQFQNGRCTLDGLVGTTPVSLSHVAKIRGTSNTV-- 298  
 QY 297 ENHVMNQVTNINGTPTDPTGDPVAPLGTDFSG-----KLFGLVLSQRDHONACKSHD 349  
 DB 299 -----INLTDLTDPFHPF-EGPAPITGFPDLGCGDWHIINWTFQGHSSQTYD----- 344  
 QY 350 AVIATNSAKFTPKLGAIQ-----IG--TWEEEDVHINQPTKFTPVGLFENEGFNQW 398  
 DB 345 --VDTPDTSPVHLSQANGISGNVIGVLSWSVPSH-----PSG-----SQVDLW 390  
 QY 399 TLPNYSGALTINMGLAPPV-APTFPGEQILPFRSHIPLKGVADPVIDCLLPQEWIOHLY 457  
 DB 391 KIPNYGSSITEATHLAPSVSGF-GEVLVFFMSKIPGPGGDSLP---CLLPQGYISHLA 446  
 QY 458 QESAPSQSDVALIRFTNPDGTGRVLFPEAKLHRSVITVANTG--SRPIVWPANGYFRPDTW 515  
 DB 447 SEQAPTVEGEPPLHLYVDPTDRNLGEFAKYPDGLTICVPNGASSGPPQLPINGVVFVSW 506  
 QY 516 VNOFYSLAPMGNGRR 532  
 DB 507 VSRFQLKEVGTASTAR 523

RESULT 15

AAB49703

ID AAB49703 standard; protein; 544 AA.

XX AAB49703;

XX AAB49703;

XX AAB49703;

DT 04-APR-2001 (first entry)

DE Small round structured virus protein SEQ ID 4.

XX Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.

XX Small round structured virus.

XX Small round structured virus.

XX Small round structured virus.

XX Small round structured virus.

PN WO200079280-A1.

XX 28-DEC-2000.  
 PD  
 XX 22-JUN-2000; 2000WO-JP004095.  
 PF  
 XX 22-JUN-1999; 99JP-00175928.  
 PR  
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 PA (DENK-) DENKA SEIKEN KK.  
 XX Takeda N, Natori K, Miyamura T, Kanata K, Sato T, Sato S;  
 XX WPI: 2001-080848/09.  
 XX N-PSDB; AAF29144.  
 DR  
 XX Kit for the detection and typing of small round-structured virus (SRSV)  
 XX strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.  
 PT  
 XX Claim 1; Page 47-49; 84pp; Japanese.  
 XX This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 CC  
 XX Sequence 544 AA;  
 Query Match 40.3%; Score 1145; DB 4; Length 544;  
 Best Local Similarity 45.1%; Pred. No. 4.6e-96;  
 Matches 249; Conservative 95; Mismatches 178; Indels 30; Gaps 14;

QY 1 MKMASNDAAAPSDGAG---LVPEAN-NETMALEPVAGASIAAPLTGQNNIIDPWIRLNF 56  
 DB 1 MMASKDATSSVDGASASQVLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINF 60  
 QY 57 VQAPNGEFTVSPRNSPCEVLLNLELGPENLPYLAHLRMYNGYAGGVEQVLLAGNAFTA 116  
 DB 61 VQAPNGEFTISPNTTSGDVLFDLSLGFHLNPFLLHLSQYNGWGMVRVIMLAGNAFTA 120  
 QY 117 GKLVFAAVPPHFFLENIISPGQITMFPVHVIDVRLTEPVLPLPDVNRNFFHYNQNEPRM 176  
 DB 121 GKLIIVCIPFGSGQQLTIAQATLFPVHVIDVRLTEPVLPLPDVNRNFFHYNQNEPRM 179  
 QY 177 RLVMAMLYTPURSGSGDDVFTVSCRVLTRPSDFDNLYVPPTLESKTKPFTLPILT 234  
 DB 180 RLVCMLYTPURSGSGDDVFTVSCRVLTRPSDFDNLYVPPTLESKTKPFTLPILT 239  
 QY 235 LTNRSRFPVPIDELYTSPNESLVVQNGRCALDGLQGTTLQTLTAICFRGRINOKV 294  
 DB 240 KYLSNRIPNPIEGMSLSPDQNTQVQFQNGRCTLDGLVGTTPVSLSHVAKIRGTSNTV--T 296  
 QY 295 SGENHVMNQVTNINGTPTDPTGDPVAPLGTDFSGKLFGLVLSQRDHONACKSHDAY--- 351  
 DB 297 SGQR---VLNLTDLGSPFAFA-APAPAGPDLGSCDWHIEMSKIPNSSTQNNPVTNS 352  
 QY 352 IATNSAKFTPKLGAIQIGTWEEDDV-----HINQTKFTPVGLFENEGNOMTLPNYSGA 406  
 DB 353 VKPNSQOFVPHLSSITL-----DENVSSGGDYGITQWTSPPSDSGGANTNFKPIDYGS 408  
 QY 407 LTNMGLAPPVAPTFPGEQILPFRSHIPLKGVGADP-VIDCLLPQEWIOHLYOESAPSQS 465  
 DB 409 LAEASQAPAVYPPGFNEVIVPMASIPGPNQSGSNLVPCLLPQGYISHLA 468  
 QY 466 DVALIRFTNPDGTGRVLFPEAKLHRSVIT-VANTGSR-PIVWPANGYFRDPTWNOFYSLA 523  
 DB 469 EALLHVVDPDTNRNLGEFAKYPGGLYTCVPNSSSTGPPQLPDGLGVFVFWVSRFYQLK 528  
 QY 524 PMGT-NGRRRV 534  
 DB 507 VSRFQLKEVGTASTAR 523



Db 529 PVGTAGPARGRL 540

Search completed: June 1, 2004, 13:45:58  
Job time : 47.9597 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6.  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.7466 Seconds  
(without alignments)  
2166.837 Million cell updates/sec

Title: US-09-926-799-8  
Perfect score: 2843  
Sequence: 1 MMASNDAPSDGAGLVP.....VNOFYSLAPMGTCGRRRVQ 535

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1196.5	42.1	530	4	US-08-486-049-3
2	291	10.2	622	4	US-09-590-020-6
3	290	10.2	623	4	US-09-590-020-4
4	286	10.1	623	4	US-09-590-020-2
5	284	10.0	668	4	US-09-617-594A-4
6	277.5	9.8	669	4	US-09-617-594A-2
7	273.5	9.6	626	4	US-09-590-020-7
8	122	4.3	980	2	US-08-473-553A-6
9	122	4.3	985	2	US-08-473-553A-2
10	121.5	4.3	2318	3	US-09-091-219-24
11	121.5	4.3	2318	4	US-09-660-541-24
12	113	4.0	2206	1	US-07-852-260-2
13	113	4.0	2206	2	US-08-461-503-2
14	113	4.0	2206	3	US-08-465-250-2
15	111	3.9	761	2	US-08-124-981A-2
16	110	3.9	716	1	US-08-396-479B-4
17	110	3.9	716	1	US-08-818-823-4
18	110	3.9	716	3	US-09-037-190-38
19	110	3.9	716	3	US-09-037-192-38
20	110	3.9	716	3	US-09-037-143-38
21	110	3.9	716	3	US-09-049-691-38
22	110	3.9	716	3	US-08-260-174-38
23	110	3.9	716	4	US-09-338-128A-38
24	110	3.9	716	4	US-09-232-346-38
25	110	3.9	716	4	US-09-037-192-38
26	109	3.8	716	5	PCT-US94-07297-37
27	109	3.8	913	4	US-09-623-624-2

28	107	3.8	716	3	US-09-037-190-46	Sequence 46, Appl
29	107	3.8	716	3	US-09-037-192-46	Sequence 46, Appl
30	107	3.8	716	3	US-09-037-143-46	Sequence 46, Appl
31	107	3.8	716	3	US-09-049-691-46	Sequence 46, Appl
32	107	3.8	716	3	US-08-260-174-46	Sequence 46, Appl
33	107	3.8	716	4	US-09-338-128A-46	Sequence 46, Appl
34	107	3.8	716	4	US-09-232-346-46	Sequence 46, Appl
35	107	3.8	716	4	US-09-037-192-46	Sequence 46, Appl
36	103	3.6	1560	4	US-09-264-512B-2	Sequence 2, Appl
37	101	3.6	2736	4	US-09-252-991A-30227	Sequence 30227, A
38	100.5	3.5	1091	6	5516630-2	Patent No. 5516630
39	100.5	3.5	2227	3	US-08-475-886-2	Sequence 2, Appl
40	100.5	3.5	2227	3	US-08-475-886-4	Sequence 4, Appl
41	100.5	3.5	2227	3	US-08-475-886-6	Sequence 6, Appl
42	100.5	3.5	2227	3	US-08-397-232-2	Sequence 2, Appl
43	100.5	3.5	2227	3	US-08-397-232-4	Sequence 4, Appl
44	100.5	3.5	2227	3	US-09-171-387-2	Sequence 2, Appl
45	100.5	3.5	2227	4	US-09-653-499-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1  
US-08-486-049-3  
; Sequence 3, Application US/08486049  
; Patent No. 6572862  
; GENERAL INFORMATION:  
; APPLICANT: Bates, Mary K  
; APPLICANT: Jiang, Xi  
; APPLICANT: Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: Fulbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE:  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-486-049-3

Query Match 42.1%; Score 1196.5; DB 4; Length 530;  
Best Local Similarity 46.4%; Pred. No. 2.1e-108;  
Matches 254; Conservative 81; Mismatches 172; Indels 41; Gaps 10;  
QY 1 MMASNDAPSDGAG---LVPEAN-NETMALEPVAGASIAAPLTGQNIIDPWIRLNF 56  
DB 1 MMASKDATSSVDGASGAGQLVPEYNASDPLAMDVPAGSSTAVATAGQVNPIDPWIRNF 60

QY 57 VQAPNGEFTVSPRNSPGVLLNLELGPBLNLYLAHLRMVNGYAGGVEVQVLLAGNAFTA 116  
DB 61 VQAPQGEFTISPNNTPGDVLFDLSLGPBLNPLFLHLQMYNGWGNRVRIMLAGNAFTA 120  
QY 117 GKLVFAANVPPHPLNIPSGQITMPPHVIIDVRLTLEPVLPLPDVRNNFFHYNQNEPRM 176  
DB 121 GKIIIVSCIPPGGSHNLTAQATLPHVIAVADTLDPIEVLDPVRLVFNHNDNQOTM 180  
QY 177 RLVAMLYTPLRSGSGDDVFTVSCRVLTFRSPDPDFNVLPTLESKTKPTLILIGE 236  
DB 181 RLVCMLYTLRTGGTGDGDFVAGRVMTCPSPDFNLEFLVPTVEQKTRPTLNLPLSS 240  
QY 237 LNSRPPVDELITYSPNLSLVQONGRCALDGELOQTOLLPTAICSPFRGRINQKVS 296  
DB 241 LNSRAPLPISSMGISPDNVQVQFQNGRCCTLDGLRGVGTTPVLSHVAKIRGTSNGTV-- 298  
QY 297 ENHVMNQVNTNGTPTDPTGDPAPLGTDFSG-----KLFGLVQRDHNAACRSHD 349  
DB 299 -----INLTEDGTFFHPF-BGAPAPIGFDLGGCDWHINMTQFGHSSQTQYD----- 344  
QY 350 AVIATNSAKTPKGLAIQ--IGTWEDDVHINQPTKFTPVGLFENEGFNOWTLPNYSGA 406  
DB 345 --VDTPTDTPVPLHLSIQANGISGN-----YGVLSWISPPHPSGSDVLDKIPNYGSS 398  
QY 407 LTLNMGAPVAPTFPGEQILFRSHIPLKGGVADPVIDCLLPQEWIQLHQSAPSD 466  
DB 399 ITEATHLAPSVPPGGEVLVFFWMSKMPGAYNLP---CLLPQEWIQLHQSAPSD 455  
QY 467 VALIRFTNPTDGRVLFPAKLRHSGYITVANTG--SRPIVVPANGYFREDTWNQVSLAP 524  
DB 456 AALLHYVDPDTRNLGFEKAYPDGFLTCVPNGASSGQQLPFGVFFVSWSRFQLKPE 515  
QY 525 MGTGNGR 532  
DB 516 VGTASSAR 523

RESULT 2  
US-09-590-020-6  
; Sequence 6, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 622  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-6

Query Match 10.2%; Score 291; DB 4; Length 622;  
Best Local Similarity 23.6%; Pred. No. 1.6e-19;  
Matches 127; Conservative 77; Mismatches 190; Indels 144; Gaps 25;  
QY 13 DGAAGLVPEANNETMA-----LEPVAGASTAAPLTQNNIIDPW-IRLNFVQAPNGEFTVS 67  
DB 79 EGGDGSITTEPQGTWVGVIAPESQAQMSAADMATGKSVDSWEAFFSHTSVNM-----S 134  
QY 68 PRNSPGEVLNLELGPBLNLYLAHLRMVNGYAGGVEVQVLLAGNAFTAAGLVFAAVPPH 127  
DB 135 TSETQKILFKQSLGPLNPLSHLAKLYVALAGSVVEFRSISGSGVFGGKLAIVVPP- 193  
QY 128 FPLENISPGQITMPPHVIIDVRLTLEPVLPLPDVRNNFFHYNQNEPRMRLVAMLYTPLR 187  
DB 194 -GEPVQSTSMQLQYPHVLFDAQVEVIFAIPLDRSLNLYHLMSTD-TTSLVIMVYNDL- 250  
QY 188 SNGSGDDVFTVSC--RVLTFRSPDPDFNVLPTLESKTKPTLILITIGELTNSRFP-- 243  
DB 251 INFVANDTSSGCIIVTETKPGDPFKPHLLKPPG-----SMLTSGSVPSDLIPKS 300  
QY 244 --VPIDELITYSPNESLVVQF---QNGRCALDGELOQTOLLPTAICSPFRGRINQKVS 298

QY 128 FPLENISPGQITMPPHVIIDVRLTLEPVLPLPDVRNNFFHYNQNEPRMRLVAMLYTPLR 187  
DB 194 -GIEPVQSTSMQLQYPHVLFDAQVEVIFAIPLDRSLNLYHLMSTD-TTSLVIMVYNDL- 250  
QY 188 SNGSGDDVFTVSC--RVLTFRSPDPDFNVLPTLESKTKPTLILITIGELTNSRFP-- 243  
DB 251 INFVANDTSSGCIIVTETKPGDPFKPHLLKPPG-----SMLTSGSVPSDLIPKS 300  
QY 244 --VPIDELITYSPNESLVVQF---QNGRCALDGELOQTOLLPTAICSPFRGRINQKVS 298  
DB 301 SSLIWIHNRHSDITFIIRPFVQNR-----TPFR-----PITIVSES 353  
QY 299 HYVMNQVNTNGTPTDPTGDPAPLGTDFSGKLVQORDHNAACRSHDAVATNSAK 358  
DB 338 --WS-----TPFR-----PITIVSES 353  
QY 359 FTKLGAIGT-----WEEDDVHINQPTKFTPVGLFENEGFNOWTLPNYSG 405  
DB 354 NMSKLG-IGVATDIYVPGIPDGMPTTI-----PEQITPAGIYSITASNGDITTAAGYDA 408  
QY 406 ALTL-----NMGAPPVAPTFPGEQI--LFRSHIPLKGGVADP--VID---CLLPQE 451  
DB 409 AETIVNTFKSMYICSLQRAWDKISNTAFITTAVRKGNISBPSNTIDMTKLIVVYOD 468  
QY 452 WIOHLYQESAPSDVALIRFT-----NPDTRVLFPAKLRHSGYITVANTGSRPI 502  
DB 469 --AHVGEVQTSITLALLGYTGIGBEAIGSDORDKVVRSVLGETG---ARGNHPI 520

RESULT 3  
US-09-590-020-4  
; Sequence 4, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-4

Query Match 10.2%; Score 290; DB 4; Length 623;  
Best Local Similarity 23.4%; Pred. No. 2e-19;  
Matches 126; Conservative 79; Mismatches 189; Indels 144; Gaps 25;  
QY 13 DGAAGLVPEANNETMA-----LEPVAGASTAAPLTQNNIIDPW-IRLNFVQAPNGEFTVS 67  
DB 79 EGGDGSITTEPQGTWVGVIAPESQAQMSAADMATGKSVDSWEAFFSHTSVNM-----S 134  
QY 68 PRNSPGEVLNLELGPBLNLYLAHLRMVNGYAGGVEVQVLLAGNAFTAAGLVFAAVPPH 127  
DB 135 TSETQKILFKQSLGPLNPLSHLAKLYVALAGSVVEFRSISGSGVFGGKLAIVVPP- 193  
QY 128 FPLENISPGQITMPPHVIIDVRLTLEPVLPLPDVRNNFFHYNQNEPRMRLVAMLYTPLR 187  
DB 194 -GEPVQSTSMQLQYPHVLFDAQVEVIFAIPLDRSLNLYHLMSTD-TTSLVIMVYNDL- 250  
QY 188 SNGSGDDVFTVSC--RVLTFRSPDPDFNVLPTLESKTKPTLILITIGELTNSRFP-- 243  
DB 251 INFVANDTSSGCIIVTETKPGDPFKPHLLKPPG-----SMLTSGSVPSDLIPKS 300  
QY 244 --VPIDELITYSPNESLVVQF---QNGRCALDGELOQTOLLPTAICSPFRGRINQKVS 298

Db 301 SLLWGNRHSWSDITFIIRPFVQNR-----HFDNQETAG-- 337  
QY 299 HYNNQVNTNNGTPDPTGDPAPLGTGDFSGKLFGLVLSQRDHDNACRSHDAVIATNSAK 358  
Db 338 --WS-----TPRFR-----PITITVSES 353  
QY 359 FTKPLGATQIGT-----WEEDDVHINOPTKFTPVGLFE-----NEGFWNTLPTNSG 405  
Db 354 NMSKLG-IGVATDIYVPGIPDGPDPDTTI-----PEQLTPAGIYSITASNGTDITTAAGYDA 408  
QY 406 ALTL-----NMGLAPPVAPTFPGEQI--LFFRSHIPKGGVADP--VID--CLLPQE 451  
Db 409 AETIVNTNFKSMYICGSLQRAWGDKISNTAFITAVRKGNISIEPSNTIDMTKLVVYQD 468  
QY 452 WIOHLYQESAPSQSDVALIRFT-----NPDTRGVLFPEAKLHRSGYITVANTGSRPI 502  
Db 469 --AHVGEVQTSIDITALLGYTGIEEAGISDRDKVVRISVLPTG-----ARGGNHPI 520  
RESULT 4  
US-09-590-020-2  
; Sequence 2, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-2

Query Match 10.1%; Score 286; DB 4; Length 623;  
Best Local Similarity 23.4%; Pred. No. 4.9e-19;  
Matches 126; Conservative 78; Mismatches 190; Indels 144; Gaps 25;  
QY 13 DGAAGLVEANNETWA----LEPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVS 67  
Db 79 EGGDGSITTEQGTWVGVIAPSPAQMSAAADMATGKSVDSWEAFPSFTSVNW----S 134  
QY 68 PRNSPGEVLLNLELGPENLPYLAHLSRMVYAGVGVQVLLAGNAFTAGKLVFAAVPPH 127  
Db 135 TSETQGLKGLPLNLPYLHAKLVAVWSGVSEVRFSGVGGKLAIVVPP- 193  
QY 128 PLENISPGQITMFFHVIDVRLTLEPVLPLPDVRNFFHYNQNEPRMLVAMLYTFLR 187  
Db 194 -GIEPVQSTSMQLQYFHLVFDARQVEVIFAIPLDRLSNLYHLSMTD-TTSLVIMVYNDL- 250  
QY 188 SNGSGDDVFTVSC--RVLTRSPDFDNVLPPTLESKTKPTLPIITIGELTNSRFP-- 243  
Db 251 INPYANDNTSSGCIIVTETKPGDFKPHLLKPPG-----SMLTGSGVPSDLIPKS 300  
QY 244 --VPIDELYTPSNESLVQVQ--QNGRCALDGELOQTTLPTAICSFRGRINQKVSSEN 298  
Db 301 SLLWGNRHSWSDITFIIRPFVQNR-----HFDNQETAG-- 337  
QY 299 HYNNQVNTNNGTPDPTGDPAPLGTGDFSGKLFGLVLSQRDHDNACRSHDAVIATNSAK 358  
Db 338 --WS-----TPRFR-----PITITVSES 353  
QY 359 FTKPLGATQIGT-----WEEDDVHINOPTKFTPVGLFE-----NEGFWNTLPTNSG 405

Db 354 NMSKLG-IGVATDIYVPGIPDGPDPDTTI-----PEQLTPAGIYSITASNGTVITTAAGYDA 408  
QY 406 ALTL-----NMGLAPPVAPTFPGEQI--LFFRSHIPKGGVADP--VID--CLLPQE 451  
Db 409 AETIVNTNFKSMYICGSLQRAWGDKISNTAFITAVRKGNISIEPSNTIDMTKLVVYQD 468  
QY 452 WIOHLYQESAPSQSDVALIRFT-----NPDTRGVLFPEAKLHRSGYITVANTGSRPI 502  
Db 469 --AHVGEVQTSIDITALLGYTGIEEAGISDRDKVVRISVLPTG-----ARGGNHPI 520  
RESULT 5  
US-09-617-594A-4  
; Sequence 4, Application US/09617594A  
; Patent No. 6541458  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.1  
; CURRENT APPLICATION NUMBER: US/09/617,594A  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 4  
; LENGTH: 668  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-617-594A-4  
Query Match 10.0%; Score 284; DB 4; Length 668;  
Best Local Similarity 26.2%; Pred. No. 8.6e-19;  
Matches 110; Conservative 49; Mismatches 125; Indels 136; Gaps 21;  
QY 17 GLVPEANNETWALEPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVS PRNSPGEV 75  
Db 141 GVIAEPNAQMSAVADVA-----TGK-SVDSEWEAFPSFTSVNW----STSETQSKI 187  
QY 76 LLNLELGPENLPYLAHLSRMVYAGVGVQVLLAGNAFTAGKLVFAAVPPHFPLENIS 135  
Db 188 LFKQSLGPLNLPYLHAKLVAVWSGVSEVRFSGVGGKLAIVVPP-----GIDP 242  
QY 136 GQIT---MFPHVVIDVRLTLEPVLPLPDVRNFFHYNQNEPRMLVAMLYTPL----- 186  
Db 243 VQSTSMQLQYFHLVFDARQVEVIFITPDLRNSLYHLSMTD-TTSLVIMVYNDLINPYAN 301  
QY 187 RNSGSGDDVFTVSCRVLTSPDFDNVLPPTLESKTKPTLPIITIGELTNSRFP-- 243  
Db 302 DSNSSGCIIV--TVETKPGDFKPHLLKPPG-----SMLTGSGVPSDLIPKS 346  
QY 244 -VPIDELYTPSNESLVQVQ--QNGRCALDGELOQTTLPTAICSFRGRINQKVSSEN 299  
Db 347 SLWIGNRHSWSDITDVIKPFVQNR-----HFDNQETAG-- 382  
QY 300 VNNQVNTNNGTPDPTGDPAPLGTGDFSGKLFGLVLSQRDHDNACRSHDAVIATNSAK 359  
Db 383 --WS-----TPRFR-----PITITVSEK 399  
QY 360 TPKLGAIOIGT-----WEEDDVHINOPTKFTPVG--LFENEGFNO-WTLPNYSGA 406  
Db 400 GSKLG-IGVATDSIVPGIPDGPDPDTTI-----PEKLTAPAGYAITNGGNDITTAADYDGA 454  
RESULT 6  
US-09-617-594A-2  
; Sequence 2, Application US/09617594A  
; Patent No. 6541458  
; GENERAL INFORMATION:

APPLICANT: Audonnet, et al.  
TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V  
FILE REFERENCE: 454313-3151.1  
CURRENT APPLICATION NUMBER: US/09/617,594A  
CURRENT FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/193,332  
PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: France 00 01761  
PRIOR FILING DATE: 2000-02-11  
PRIOR APPLICATION NUMBER: France 99 09421  
PRIOR FILING DATE: 1999-07-16  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 2  
LENGTH: 669  
TYPE: PRT  
ORGANISM: Feline calicivirus  
US-09-617-594A-2

Query Match 9.8%; Score 277.5; DB 4; Length 669;  
Best Local Similarity 28.4%; Pred. No. 3.7e-18;  
Matches 103; Conservative 49; Mismatches 152; Indels 59; Gaps 16;

QY 30 EPVAGASIAAPLTGQNNIIDPWIR-LNFVQAPNGEFTVSPRNSPGEVLLNLELGPENPY 88  
DB 146 EPSAQMATADAATGKSDSEWSEFESFHTSVNM-----STSETQKILFKQSLGPELLNPY 201  
QY 89 LAHLSRMVYAGGVEVQVLLAGNAFTAGKLVFAAVPPHPFPLENISPGQIT---MFPHVI 145  
DB 202 LEHLSKLYVAVSGSDVRFSGVFGKLAIVVPP-----GVDPPVQSTMLQYPHVL 256  
QY 146 IDVRLTEPVLPLPDVRRNFFHYNQNEPRMLVAMLYTPL-----RNSGSDDDVFTVS 199  
DB 257 FDARQVEPVIFSLDRLSTLYHLMSTD-TTSLVIMVYNDLINPYANDNSGCIY---- 311  
QY 200 CRVLTRSPDPDFNLYVPPTLESKTKPFTLPILTIGELTNSRPP-----VPIDELYSNPE 255  
DB 312 -TVETKPGDPFKHLLKPPG-----SMTHGSPDILPKSSSLWIGNRYWSDIT 360  
QY 256 SLVVOP---QNGRCALDGELOQ--TTQLLPATICSFRGRINQKVGSEHVMNQVTNI- 308  
DB 361 DFVIRPFVQNRHFDNQETAGWSTFRFPITIT--SESNGSKLGTGVATDIYVGP 418  
QY 309 NGTFPDPDTPGVPAPLGTPTDPSGKLVGLVSRDHD-----NACRSHDAVIATNSAKFTKLG 364  
DB 419 DGWPDTTIGELTBAG--DYS-----ITNGSGNDIATANAYDSADVITNTNFRGMVIG 471  
QY 365 AIQ 367  
DB 472 ALQ 474

RESULT 7  
US-09-590-020-7  
Sequence 7, Application US/09590020  
Patent No. 6355246  
GENERAL INFORMATION:  
APPLICANT: Kruger, John M  
APPLICANT: Maes, Roger K  
APPLICANT: Vilnis, Alvars  
TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
FILE REFERENCE: MSU 4.1-446  
CURRENT APPLICATION NUMBER: US/09/590,020  
CURRENT FILING DATE: 2000-06-08  
PRIOR APPLICATION NUMBER: 60/138,484  
PRIOR FILING DATE: 1999-06-10  
NUMBER OF SEQ ID NOS: 9  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 7  
LENGTH: 626  
TYPE: PRT  
ORGANISM: Feline calicivirus

US-09-590-020-7

Query Match 9.6%; Score 273.5; DB 4; Length 626;  
Best Local Similarity 32.3%; Pred. No. 8.3e-18;  
Matches 71; Conservative 36; Mismatches 86; Indels 27; Gaps 8;

QY 30 EPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPGEVLLNLELGPENPY 88  
DB 100 EPSAQMATADAATGKSDSEWSEFESFHTSVNM-----STSETQKILFKQSLGPELLNPY 155  
QY 89 LAHLSRMVYAGGVEVQVLLAGNAFTAGKLVFAAVPPHPFPLENISPGQIT---MFPHVI 145  
DB 156 LSHLAKLYVAVSGSIEVRESISGSGVFGKLAIVVPP-----GVDPPVQSTMLQYPHVL 210  
QY 146 IDVRLTEPVLPLPDVRRNFFHYNQNEPRMLVAMLYTPLRNSGSDDDVFTVSC--RYL 203  
DB 211 FDARQVEPVIFCLDRLSTLYHLMSTD-TTSLVIMVYNDLINPYANDNTSSGCIYVTE 268  
QY 204 TRPSDPDFNLYVPPTLESKTKPFTLPILTIGELTNSRPP 243  
DB 269 TKPGDPFKHLLKPPG-----SMLTHGSPVPSDLIP 298

RESULT 8

US-08-473-553A-6  
Sequence 6, Application US/08473553A  
Patent No. 5859338  
GENERAL INFORMATION:  
APPLICANT: Meyerowitz, Elliot M.  
APPLICANT: Clark, Steven E.  
APPLICANT: Williams, Robert W.  
TITLE OF INVENTION: Plant Clavatal Nucleic Acids,  
TITLE OF INVENTION: Transformed Plants, and Proteins  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hobbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: California  
COUNTRY: United States  
ZIP: 94111-4187  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/473,553A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 800  
ATTORNEY/AGENT INFORMATION:  
NAME: Silva, Robin M.  
REGISTRATION NUMBER: 38,304  
REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 781-1989  
TELEFAX: (415) 398-3249  
TELEX: 910 277299  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 980 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: protein  
US-08-473-553A-6

Query Match 4.3%; Score 122; DB 2; Length 980;  
Best Local Similarity 20.9%; Pred. No. 0.012;  
Matches 119; Conservative 64; Mismatches 157; Indels 230; Gaps 30;

QY 90 AHLSRM-----YNGYAGV-----EVQVLLAGNAFTAGKLVFAA----- 123

Db 210 AFLSLKLNREMYIGYNSYTCGVPRFGLTKLBILDWASCTLTGTEIPTSLNKLHHT 269  
QY 124 -----VPPHPPLE-----NISPGQIT-MFPHVILDVRLTLEPVLPLPVRNNF 165  
Db 270 LFLHNNLTGHIPPELSGLVSLKSLDLSINQLTGEBIPOSFINLGNITLINL-----FRNNL 325  
QY 166 FHYNQONE-----PRMLVAM-----LYTPRSNGSGDDV-FTVSCRVLTRPSPD----- 209  
Db 326 --YGOIPEAIGELPKLEVEFWENNFTLQLPANLGRNGNLKLDVSDNHLTGLIPKOLCR 383  
QY 210 -----PDFNYLVPPTLE-----SKTK-----ESLVVQPNQGRCALDDEL----- 272  
Db 384 GEKLEMLILSNFPFGPIPEELGCKSLTKIRIVKNLLNGTVPAGLFNPLVITIELTDN 443  
QY 241 RF-----PVP-----IDELTSPN-----ESLVVQPNQGRCALDDEL----- 272  
Db 444 PFGELPVTMSGDVLDQIYLSNNMFSGEIPPAIGNFNNLQTLFLDRNFRGNIPREIPEL 503  
QY 273 -----QOTTOLLPTAI--CS-----FRGRINQKV-SGENHYVNMQVTTINGTP 312  
Db 504 KHLSRINTSANNITGGIPDSISRCSTLISVDLSRNRINGEIPKGINNVKNLGLTINISNQ 563  
QY 313 FDTGDPVAPLGT-----PDFSGK-----LFGVLSQRDHDNACRSHDAVIA 353  
Db 564 L--TGSIFTGIGNMTSLTTLDSFNDLSGRVPLGQFLVFNETSFAGNTYLCPLHRVSCP 621  
QY 354 T-----NSAKFTPK-----LGAIQITWEEDDVHINQPTKFTPVGLFENE 394  
Db 622 TRPGQTSDDHNTALFSPSRIVITVIAAITGLILI-----SVAIRQMNK-----KKQK 669  
QY 395 FNQWTLPNYSGALTINMGLAPPVAPTFPGEQILFPRSHIPLKGGVADPVIDCLLPQEWIQ 454  
Db 670 SLAWKLTAFA-----QKLDKFS-----EDVLECKEENIIG 699  
QY 455 H-----LYQESAPQSODVALIRFTNPDTRGR 479  
Db 700 KGGGIVYRGSMPNNVDVAIKRLVGRGTGR 729

## RESULT 9

US-08-473-553A-2  
; Sequence 2, Application US/08473553A  
; Patent No. 5859338  
; GENERAL INFORMATION:  
; APPLICANT: Meyerowitz, Elliot M.  
; APPLICANT: Clark, Steven E.  
; APPLICANT: Williams, Robert W.  
; TITLE OF INVENTION: Plant Clavatal Nucleic Acids,  
; ADDRESSSEE: Flehr, Hobbach, Test, Albritton & Herbert  
; CORRESPONDENCE ADDRESS:  
; STREET: Four Embarcadero Center, Suite 3400  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: United States  
; ZIP: 94111-4187  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/473,553A  
; FILING DATE: 06-JUN-1995  
; CLASSIFICATION: 800  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Silva, Robin M.  
; REGISTRATION NUMBER: 38,304  
; REFERENCE/DOCKET NUMBER: A-60886/RFT/RMS  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 781-1989  
; TELEFAX: (415) 398-3249

; TELEX: 910 277299  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 985 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-473-553A-2

Query Match 4.3%; Score 122; DB 2; Length 985;  
Best Local Similarity 20.9%; Pred. No. 0.012;  
Matches 119; Conservative 64; Mismatches 157; Indels 230; Gaps 30;

QY 90 AHLSSM-----YNGYAGV-----EVLVLLAGNAFTAGKLVFAA----- 123  
Db 215 AFLSLKLNREMYIGYNSYTCGVPRFGLTKLBILDWASCTLTGTEIPTSLNKLHHT 274  
QY 124 -----VPPHPPLE-----NISPGQIT-MFPHVILDVRLTLEPVLPLPVRNNF 165  
Db 275 LFLHNNLTGHIPPELSGLVSLKSLDLSINQLTGEBIPOSFINLGNITLINL-----FRNNL 330  
QY 166 FHYNQONE-----PRMLVAM-----LYTPRSNGSGDDV-FTVSCRVLTRPSPD----- 209  
Db 331 --YGOIPEAIGELPKLEVEFWENNFTLQLPANLGRNGNLKLDVSDNHLTGLIPKOLCR 388  
QY 210 -----PDFNYLVPPTLE-----SKTK-----PPTLPILTIGELTNS 240  
Db 389 GEKLEMLILSNFPFGPIPEELGCKSLTKIRIVKNLLNGTVPAGLFNPLVITIELTDN 448  
QY 241 RF-----PVP-----IDELTSPN-----ESLVVQPNQGRCALDDEL----- 272  
Db 449 PFGELPVTMSGDVLDQIYLSNNMFSGEIPPAIGNFNNLQTLFLDRNFRGNIPREIPEL 508  
QY 273 -----QOTTOLLPTAI--CS-----FRGRINQKV-SGENHYVNMQVTTINGTP 312  
Db 509 KHLSRINTSANNITGGIPDSISRCSTLISVDLSRNRINGEIPKGINNVKNLGLTINISNQ 568  
QY 313 FDTGDPVAPLGT-----PDFSGK-----LFGVLSQRDHDNACRSHDAVIA 353  
Db 569 L--TGSIFTGIGNMTSLTTLDSFNDLSGRVPLGQFLVFNETSFAGNTYLCPLHRVSCP 626  
QY 354 T-----NSAKFTPK-----LGAIQITWEEDDVHINQPTKFTPVGLFENE 394  
Db 627 TRPGQTSDDHNTALFSPSRIVITVIAAITGLILI-----SVAIRQMNK-----KKQK 674  
QY 395 FNQWTLPNYSGALTINMGLAPPVAPTFPGEQILFPRSHIPLKGGVADPVIDCLLPQEWIQ 454  
Db 675 SLAWKLTAFA-----QKLDKFS-----EDVLECKEENIIG 704  
QY 455 H-----LYQESAPQSODVALIRFTNPDTRGR 479  
Db 705 KGGGIVYRGSMPNNVDVAIKRLVGRGTGR 734

## RESULT 10

US-09-091-219-24  
; Sequence 24, Application US/09091219  
; Patent No. 6171592  
; GENERAL INFORMATION:  
; APPLICANT: STUDDERT, Michael J.  
; APPLICANT: CRABB, Brendan S.  
; APPLICANT: FENG, Li  
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS  
; FILE REFERENCE: 040268/0151  
; CURRENT APPLICATION NUMBER: US/09/091,219  
; CURRENT FILING DATE: 1998-10-05  
; EARLIER APPLICATION NUMBER: PCT/AU96/00815  
; EARLIER FILING DATE: 1996-12-18  
; EARLIER APPLICATION NUMBER: AU PN7201  
; EARLIER FILING DATE: 1995-12-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: Patent in Ver. 2.0  
; SEQ ID NO 24

```

LENGTH: 2318
TYPE: PRT
ORGANISM: Foot-and-mouth disease virus
IS-09-091-219-24

Query Match      4.3%; Score 121.5; DB 3; Length 2318;
Best Local Similarity 26.0%; Pred. No. 0.057;
Matches 38; Conservative 21; Mismatches 62; Indels 25; Gaps 2;

19y 79 LELGPELNPYLAAHLSRMVNGYAGGVEVQVLLAGNAFTAGKLVFAAYPHPHLENISPGQI 138
    ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
16b 367 LELPTDHKGVGSGLTDSAYMRNGWDVEVTAVGNQFNGGCLLVAMVPPELYSIQKRELYQL 426
    ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

19y 139 TMFPFHVIDVRLTEPLVLLPDVRNFFHYNQNEPRMRLVAMLYTPLRSNGSG----- 192
    |||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
16b 427 TLFPHQFINPRTNMTAHITVPFGVNRDYQKVHKP-WTLVVMVVAFLVTNTEGAPQIKV 485
    |||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

193 -----DDVFTVSC 200
486 YANIAPTNVHVAGEPPSKEGIFPVAC 511

RESULT 11
US-09-660-541-24
; Sequence 24, Application US/09660541
; Patent No. 6531136
; GENERAL INFORMATION:
; APPLICANT: STUDDERT, Michael J.
; APPLICANT: CRABB, Brendan S.
; APPLICANT: FENG, Li
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS
; FILE REFERENCE: 040268/0151
; CURRENT APPLICATION NUMBER: US/09/660,541
; CURRENT FILING DATE: 2000-09-12
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/091,219
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 24
; LENGTH: 2318
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-09-660-541-24

Query Match      4.3%; Score 121.5; DB 4; Length 2318;
Best Local Similarity 26.0%; Pred. No. 0.057;
Matches 38; Conservative 21; Mismatches 62; Indels 25; Gaps 2;

19y 79 LELGPELNPYLAAHLSRMVNGYAGGVEVQVLLAGNAFTAGKLVFAAYPHPHLENISPGQI 138
    ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
16b 367 LELPTDHKGVGSGLTDSAYMRNGWDVEVTAVGNQFNGGCLLVAMVPPELYSIQKRELYQL 426
    ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

19y 139 TMFPFHVIDVRLTEPLVLLPDVRNFFHYNQNEPRMRLVAMLYTPLRSNGSG----- 192
    |||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :
16b 427 TLFPHQFINPRTNMTAHITVPFGVNRDYQKVHKP-WTLVVMVVAFLVTNTEGAPQIKV 485
    |||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| :

193 -----DDVFTVSC 200
486 YANIAPTNVHVAGEPPSKEGIFPVAC 511

RESULT 12
US-07-852-260-2
; Sequence 2, Application US/07852260
; Patent No. 5525715
; GENERAL INFORMATION:
; APPLICANT: Racanello, Vincent
; APPLICANT: Tateem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA

```



```
RESULT 13
US-08-461-503-2
; Sequence 2, Application US/08461503
; Patent No. 5834302
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES
; TITLE OF INVENTION: FROM CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,503
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-503-2

Query Match 4.0%; Score 113; DB 2; Length 2206;
Best Local Similarity 20.0%; Pred. No. 0.36;
Matches 98; Conservative 54; Mismatches 175; Indels 164; Gaps 22;

QY 48 IDPWIRLNFVQAPNG-----EFTVSPRNSPGEVLLNLELGPBLNPLYLAH-----LSRMYN 97
DB 389 IDTWIPLNLESTKRNMTDMYRVTLSDSADLSQPICLSLSPAFDPRLSHTMLGEVLNYVT 448
QY 98 GYAGVEVQVLLAGNAFTAGKLVFAAVPPHPPLENISPCQITMFPFHVIDVTLPEVLIP 157
DB 449 HWAGSLKFTFLFCGSMATGKILVAYAPPG-AQPPTSKEAMGLGTHVWDLGLQSSCTMV 507
QY 158 LPDVARNPFHYNQNE-----PRMLVAMLYTPLRNSGDDVFTVSCRVLTRPS 207
DB 508 VPWISNVYRQTQDSFTGGYISMFYQTRIVVPLSTPKMSMLG---FVSACN-----558
QY 208 PDFDFNYLVPPTLESKTKPFTLPILTIGELNRSFPVPIDELYTSPNESLVVQPNQGRCA 267
DB 559 -DFSVRLLRDTT-----HISQALPQGIEDL-----TSEVAQG-----A 591
QY 268 LDGELQGTQLLPTALCSFRGRINOKVSGENHVMNQVNTNGTFFDPTGDPVA-----321
DB 592 LTLSLPKQODSLP-----DTKASGPAH-----SKEVPALTAVET 625
QY 322 ----PLGTPDFSGKLVLSQDHD-----NACRSHDAVIATNSAKFTPKLGAIQ 367
DB 626 GATNPL-APSDVTQTRHVQRSRSESTIESFARGAC---VAIIEDVNEQPTTQAQKU-. 680
QY 368 ICTWE---EDDVHINOPTKFTVPGLFENEGFNQWTLPNYSGALTLNMGLA-----414

RESULT 14
US-08-465-250-2
; Sequence 2, Application US/08465250
; Patent No. 6136570
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM
; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release 1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,250
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-250-2

Query Match 4.0%; Score 113; DB 3; Length 2206;
Best Local Similarity 20.0%; Pred. No. 0.36;
Matches 98; Conservative 54; Mismatches 175; Indels 164; Gaps 22;

QY 48 IDPWIRLNFVQAPNG-----EFTVSPRNSPGEVLLNLELGPBLNPLYLAH-----LSRMYN 97
DB 389 IDTWIPLNLESTKRNMTDMYRVTLSDSADLSQPICLSLSPAFDPRLSHTMLGEVLNYVT 448
QY 98 GYAGVEVQVLLAGNAFTAGKLVFAAVPPHPPLENISPCQITMFPFHVIDVTLPEVLIP 157
DB 449 HWAGSLKFTFLFCGSMATGKILVAYAPPG-AQPPTSKEAMGLGTHVWDLGLQSSCTMV 507
QY 158 LPDVARNPFHYNQNE-----PRMLVAMLYTPLRNSGDDVFTVSCRVLTRPS 207
DB 508 VPWISNVYRQTQDSFTGGYISMFYQTRIVVPLSTPKMSMLG---FVSACN-----558
QY 208 PDFDFNYLVPPTLESKTKPFTLPILTIGELNRSFPVPIDELYTSPNESLVVQPNQGRCA 267
```

Db 559 -DFSRLRLDIT-----HISQSALPOGIEDL-----TSEVAQG-----A 591  
QY 268 LDGELOQTQLLPAICFRGRINQKUSGENHVMNQVTNPGFPDPTGDVPA----- 321  
Db 592 LTLSPKQDLSL-----DTKASGPAH-----SKEVPALTAVET 625  
QY 322 ----PLGTPDFSGKLFGLVSORDHD-----NACRSHDAVIATNSAKFTPKLGAIQ 367  
Db 626 GATNPL-APSDTVQRHVQRSRSESTIESFFARGAC---VAIEVDNEQPTTTRAKL- 680  
QY 368 ICTWE---EDDVHINOPTKFTPVGLFENEGNQWTLPHYSGALTINMCLA----- 414  
Db 681 FAWWRITYKDTVLQARKLEFFTYSRFDMF-FTFVVYTAFTNA---NNGHALNQVQIMYI 736  
QY 415 PPVAPTFPEQILFRSHPLKGGVADPVIDCLLPQEWIQLHYQES-----APS 463  
Db 737 PFGAPT-----PKSWDDYTWQTSNPSNPSIFYYTGAAPA 768  
QY 464 QSDVALIRFTN 474  
Db 769 RISVPVYGLAN 779

RESULT 15  
US-08-124-981A-2  
; Sequence 2, Application US/08124981A  
; Patent No. 5837840  
; GENERAL INFORMATION:  
; APPLICANT: Crabtree, Gerald R.  
; TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES  
; NUMBER OF SEQUENCES: 39  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Kourie and Crew  
; STREET: 379 Lytton Avenue  
; CITY: Palo Alto  
; STATE: California  
; COUNTRY: US  
; ZIP: 94301  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/124,981A  
; FILING DATE: 20-SEP-1993  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Smith, William M  
; REGISTRATION NUMBER: 30,223  
; REFERENCE/DOCKET NUMBER: 5490A-226  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 326-2400  
; TELEFAX: (415) 326-2422  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 761 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-124-981A-2

Query Match 3.9%; Score 111; DB 2; Length 761;  
Best Local Similarity 21.5%; Pred. No. 0.097;  
Matches 130; Conservative 66; Mismatches 235; Indels 174; Gaps 32;  
QY 14 GAAGLVPEA--NNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAPNGEFTVSPRNS 71  
Db 12 GARAVFPETPRQRGAATSPVPP--AAPLPAA-----AARMPSTSPFPV-PSKF 57  
QY 72 PGEVLLNLELGPENLPYLAHLSRMVYAGGVQVLLAGNAFTAGKLVFAAVPPH--FP 129

Db 58 P-----LGP-----AAAVFGRGETLGPAPRAGTGMKSABEEHYGYA 93  
QY 130 LENTSPGOITMFPKHII-----DVRTLEPVLLPLPDVRRNF-----FHYN 169  
Db 94 SNNVSPALPLPTAHSTLPAPCHNLQSTPGIIPADHPSGVGAALDGGPAGYFLSSGHTR 153  
QY 170 QQ-----NEPRMLVAMLYTFLRNSGDDVFTVSCVLTFRPSDFDFNVLPVPTLESKT 224  
Db 154 PDGAPALSPRIETISCLGLYHNNNQFHDVEVEDVLPSSKRSP--STATLSLSLEAYR 211  
QY 225 KPFTL-PILTIGELT-----NSRFPVPIDELVTSNESLWVQPN-----G 264  
Db 212 DPCSLSPASSLSRSCNSEASSYESINYPASPTSPWQSPCVSPKTTDPEEGFPRLGL 271  
QY 265 RCALDGLQGTQTLPTAICSPFRINQKUSGENHVMNQVTNNG---TPDTP-TGDVPA 321  
Db 272 ACTLLGSPQHSPTSPRASVTEESWLGARSPPCNKRKYSLNGRQPPYSPHHSPTPS 331  
QY 322 PLGTPDFS---GKLFQVLSQORDHACRSHDAVIATNSAKFTPKLGAIQIGTWEEDDVHI 378  
Db 332 PHGSPRVSVTDDSWLGNITQ-----YTSSAIVAAINALTTD--SSLDLG-----DGVVP 378  
QY 379 -----NOP-----TKFTPVGLFENEGNQWTLPHYSGALTINMGLAPPVAPTFPEQIL 427  
Db 379 KSRKTILEQPPSVALKVEPVG-----EDLGSPPPPADFAPEDYSS 418  
QY 428 FFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPS-----QSDVALIRFTN 474  
Db 419 F--QHI-RKGGFCQYL--AVP-----QHPYQWAKPKPLSPTSYMSPTLPALDQWQLPSHSG 469  
QY 475 PDTGRVLFQAKL-HRSGYIT-----VANTGSRPIVVPANGYFRFDTWVNOFYSLAPMG 526  
Db 470 PYELRIEVQPKSHHRAHYETEGSRGAVKASAGGHPV-VQLHGYLENEPLMLQLF-----IG 524  
QY 527 TGNGR 531  
Db 525 TADDR 529

Search completed: June 1, 2004, 13:58:04  
Job time : 14.7466 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.2005 Seconds  
(without alignments)  
4368.312 Million cell updates/sec

Title: US-09-926-799-8

Perfect score: 2843

Sequence: 1 MKMASNDAPSDGAGLVP.....VNQFSLAPMTGNGRRRVQ 535

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1196.5	42.1	530	14	US-10-314-739-3
2	284	10.0	668	14	Sequence 3, Appli
3	277.5	9.8	547	12	Sequence 4, Appli
4	277.5	9.8	669	14	Sequence 4, Appli
5	277.5	9.8	671	12	Sequence 2, Appli
6	125.5	4.4	1147	15	Sequence 2, Appli
7	117	4.1	338	10	Sequence 38, Appli
8	110.5	3.9	2468	12	Sequence 2, Appli
9	110.5	3.9	2468	14	Sequence 66335, A
10	110	3.9	716	10	US-10-246-330-4
11	110	3.9	716	14	US-09-349-058-38
12	110	3.9	716	15	US-10-040-430-38
13	109.5	3.9	925	15	US-10-116-275-182
14	109	3.8	913	14	Sequence 182, App
15	109	3.8	913	15	Sequence 3889, Ap
					Sequence 2, Appli
					Sequence 132, App

16	107	3.8	716	10	US-09-349-058-46	Sequence 46, Appli
17	107	3.8	716	14	US-10-040-430-46	Sequence 46, Appli
18	106.5	3.7	951	14	US-10-024-298A-172	Sequence 172, App
19	106.5	3.7	951	14	US-10-042-211A-172	Sequence 172, App
20	106.5	3.7	951	16	US-10-617-217A-172	Sequence 172, App
21	106.5	3.7	1592	12	US-10-231-956A-319	Sequence 319, App
22	105.5	3.7	988	12	US-10-424-599-144895	Sequence 144895,
23	104.5	3.7	5636	14	US-10-032-189-128	Sequence 128, App
24	104.5	3.7	5636	15	US-10-120-801-72	Sequence 72, Appli
25	104.5	3.7	5636	15	US-10-023-634-93	Sequence 93, Appli
26	102.5	3.6	548	11	US-09-764-875-919	Sequence 919, App
27	102.5	3.6	548	15	US-10-104-047-2916	Sequence 2916, Ap
28	102.5	3.6	3930	12	US-10-282-122A-46817	Sequence 46817, A
29	102	3.6	643	14	US-10-156-761-12113	Sequence 12113, A
30	102	3.6	2572	14	US-10-114-153-86	Sequence 86, Appli
31	102	3.6	2673	15	US-10-120-801-76	Sequence 76, Appli
32	102	3.6	3645	14	US-10-032-189-127	Sequence 127, App
33	102	3.6	3645	15	US-10-120-801-73	Sequence 73, Appli
34	102	3.6	4495	15	US-10-138-588-20	Sequence 20, Appli
35	102	3.6	5635	16	US-10-451-168-78	Sequence 78, Appli
36	101.5	3.6	1382	14	US-10-010-160-18	Sequence 18, Appli
37	100.5	3.5	836	14	US-10-272-459-40	Sequence 40, Appli
38	100.5	3.5	980	14	US-10-272-459-41	Sequence 41, Appli
39	100.5	3.5	2227	9	US-09-929-955-12	Sequence 12, Appli
40	100.5	3.5	2227	13	US-10-104-366-12	Sequence 12, Appli
41	100.5	3.5	2227	13	US-10-135-988-2	Sequence 2, Appli
42	100.5	3.5	2227	13	US-10-135-988-4	Sequence 4, Appli
43	100.5	3.5	2227	13	US-10-135-988-6	Sequence 6, Appli
44	100.5	3.5	2227	16	US-10-719-619-12	Sequence 12, Appli
45	100	3.5	1196	16	US-10-240-577-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1  
US-10-314-739-3  
; Sequence 3, Application US/10314739  
; Publication No. US20030129588A1  
; GENERAL INFORMATION:  
; APPLICANT: Bates, Mary K  
; Jiang, Xi  
; Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; Characterize No. US20030129588A1walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pubbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE: <Unknown>  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/10/314,739  
; APPLICATION NUMBER: US/10/314,739  
; FILING DATE: 09-Dec-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX: <Unknown>

```

;
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match 42.1%; Score 1196.5; DB 14; Length 530;
Best Local Similarity 46.4%; Pred. No. 3.3e-105; Indels 41; Gaps 10;
Matches 254; Conservative 81; Mismatches 172;

QY 1 MKNASNDAPSDGAG---LVPEAN-NETMALEPVAGASIAAPLTGQNNIIDPWIRLNF 56
DB 1 MMWASKDATSSVDGAGAGQLPEVNASDPLANDPVAGSTAVATAGVNPIDPWIINF 60

QY 57 VOAPNGEFTVSPRNSPEVLLNLELGPENLYLAHLSRMVNGYAGGVEQVLLAGNAFTA 116
DB 61 VQAPQGEFTISPNTGDLVFLSLGPHLNPFLHLSQMYNGWVGNMVRIMLAGNAFTA 120

QY 117 GKLVFAAVPPHPLENISQGITMPPHVIDVRLTLEPVLPLPDVRRNFFHYNQNEPRM 176
DB 121 GKIIVCIPPGFGSHNLTTAATLFPHVIAVRLTLEPVLPLPDVRRNFFHYNQNEPRM 180

QY 177 RLVMALYTPLRNSGDDVFTVSCRVLTRPSDFDFNYLVPTLESKTKPFTLPILTIGE 236
DB 181 RLVMALYTPLRNSGDDVFTVSCRVLTRPSDFDFNYLVPTLESKTKPFTLPILTIGE 240

QY 237 LTNSRPFPVDELITYSPNESLVQPNQRCALDGELOQTQLPTAICSFGRINQKVS 296
DB 241 LNSRAPLPISSMIGISPDNVQSVQFQNGRCRLDGLRVGTPVLSLHVAKIRGTSNGTV-- 298

QY 297 ENHVMNQVTNINGTRPDGDPAPLGTDFSG-----KLVGLVSRQDHDNACRSHD 349
DB 299 -----INLTEDGTTPHPH-EGPAPIGFDFDLGSCDWHINMTQFGHSSQTQYD----- 344

QY 350 AVIATNSAKFTPKLGAIQ---IGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLNYS 406
DB 345 --VDTPTDFVPHLGSIQANGIGSGN---YVGLVSWISPPSHPSGSQVDLWKIPNGS 398

QY 407 LTNMGLAPPVAPTFPGEQILFFESHPLKGSVADPVIDCLLPQEWIQLHLYQSAPSQSD 466
DB 399 ITEATHLAPSYPPGFGVLVFFKMSKMPGAYNLP---CLLPQEYISHLASEQAPTVGE 455

QY 467 VALIRFTNPDRVRLFEAKLHRSYITVANTG--SRPIVVPANGYFRFDTWVQFYSLAP 524
DB 456 AALLHYVDPTDGRNLGEPKAYPDGFLTCVPNGASSGQQLPFGVVFVSVWSRFYQLKP 515

QY 525 MGTGNRR 532
DB 516 VGTASSAR 523

RESULT 2
US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26

; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 4
; LENGTH: 668
; TYPE: PR
; ORGANISM: Feline calicivirus
; US-10-209-507-4

Query Match 10.0%; Score 284; DB 14; Length 668;
Best Local Similarity 26.2%; Pred. No. 9.6e-18; Indels 136; Gaps 21;
Matches 110; Conservative 49; Mismatches 125;

QY 17 GLVPEANETMALEPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPCEV 75
DB 141 GVIAEPNAQMSAVADVA-----TGK-SVDSEAEAFSFHTSVNM-----STSETQGI 187

QY 76 LLNLELGPENLYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKLVFAAVPPHPLENTSP 135
DB 188 LFKQSLGPLNLPYLTHLAKLYVAVSGSIEVRSISGSGVFGKLAIVVPP-----GIDP 242

QY 136 GQIT---MPPHVIDVRLTLEPVLPLPDVRRNFFHYNQNEPRMLVAMLYTPL----- 186
DB 243 VQSTSMLOYPHVLFDARQVEPVITFDLRLNSLYHLSMDT-TTSLVIMYNDLNPVAN 301

QY 187 RNSGDDVFTVSCRVLTRPSDFDFNYLVPTLESKTKPFTLPILTIGELTNSRFP--- 243
DB 302 DSNSSGCV-----TVETKPGDPFKHLLKPPG-----SMLTSGSIFSDLIPKSS 346

QY 244 -VPIDELITYSPNESLVQPN---QNGRCALDGELOQTQLPTAICSFGRINQKVS 299
DB 347 SLMIGNRHNSDIITDFVILKPFVQANR-----HFDNFQETAG--- 382

QY 300 VMNQVTNINGTRPDGDPAPLGTDFSGKLVGLVSRQDHDNACRSHDAVIATNSAKF 359
DB 383 -WS-----TPRFR-----PITITVSEK 399

QY 360 TPKLGAIQGT-----WEEDDVHINQPTKFTPVG--LFENEGFNQ-WTLNYS 406
DB 400 GSKLG-IGVATDSIVPGIPDGWPDFTI---PEKLIPAGDYAITNGNNDITTAADYDGA 454

RESULT 3
US-10-670-695-4
; Sequence 4, Application US/10670695
; Publication No. US20040059316A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Wayne A.
; APPLICANT: Lappin, Michael R.
; APPLICANT: Rosen, David K.
; APPLICANT: Andrews, Janet S.
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE
; FILE REFERENCE: DI-9-1
; CURRENT APPLICATION NUMBER: US/10/670,695
; CURRENT FILING DATE: 2003-09-25
; PRIOR APPLICATION NUMBER: 09/521,738
; PRIOR FILING DATE: 2000-03-09
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 547
; TYPE: PR
; ORGANISM: Feline calicivirus
; US-10-670-695-4

Query Match 9.8%; Score 277.5; DB 12; Length 547;
Best Local Similarity 31.4%; Pred. No. 2.9e-17; Indels 31; Gaps 9;
Matches 76; Conservative 38; Mismatches 97;

QY 30 EPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPCEVLLNLSELNPNY 88
DB 21 EPSAQMSADMATGKSVDSSEAEAFSFHTSVNM-----STSETQGIKFLQSLGPLN 76
QY 89 LAHLSRMVNGYAGGVEQVLLAGNAFTAGKLVFAAVPPHPLENTSPGQIT---MPPHVI 145
```

Db 77 LEHLAKLYVWAGSGSEVRSISGSGVFGKLAIAIVVPP-----GDPVQSTSMLOYPHVL 131  
QY 146 IDVRLPEVLLPLDVRNFFHYNQNEPRMELVAMLYTPLRSNGSGDDVFTVSC--RVL 203  
Db 132 FDARQVEPVIFCLPDLRSTLYHLMSDTD--TTSLVIMVYNDL--INPYANDANSSGGCIVTVE 189  
QY 204 TRPSDFDFNVLVPPTLESKTKPFTLPILTIGELTNSRPP-----VPIDELYTSNPSLVV 259  
Db 190 TKPGDPDFKHLKPPG-----SMLTHGSIPLDKTSSSLWIGNRYMSDITDFVI 239  
QY 260 QP 261  
Db 240 RP 241

RESULT 4  
US-10-209-507-2  
; Sequence 2, Application US/10209507  
; Publication No. US20030109033A1  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.2  
; CURRENT APPLICATION NUMBER: US/10/209,507  
; PRIOR FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: 09/617,594  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: Patentin version 3.0  
; SEQ ID NO 2  
; LENGTH: 669  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-10-209-507-2

Query Match 9.8%; Score 277.5; DB 14; Length 669;  
Best Local Similarity 28.4%; Pred. No. 4e-17;  
Matches 103; Conservative 49; Mismatches 152; Indels 59; Gaps 16;  
QY 30 EPVAGASIAAPLTGQNNIIDPW--IRLNFVQAPNGEFTVSPRNSPGEVLLNLELGPRLNPY 88  
Db 146 EPSAQMAATAADATGKSDSEWSEFSTSYNW----STSETQKILFKQSLGPLLPY 201  
QY 89 LAHLSRMVYAGGVEVQVLLAGNAFTAGKLVFAAVPPHFFPLENISPGQIT---MPPHVI 145  
Db 202 LEHLKLYVWAGSDVDRFSISGSGVFGKLAIAIVVPP-----GDPVQSTSMLOYPHVL 256  
QY 146 IDVRLPEVLLPLDVRNFFHYNQNEPRMELVAMLYTPL-----RSNGSGDDVFTVTS 199  
Db 257 FDARQVEPVIFSPDLRSTLYHLMSDTD--TTSLVIMVYNDLNPVANDSNSSGCIV---- 311  
QY 200 CRVLTRPSDFDFNVLVPPTLESKTKPFTLPILTIGELTNSRPP-----VPIDELYTSNE 255  
Db 312 -TVETKPGDPDFKHLKPPG-----SMLTHGSIPLDKTSSSLWIGNRYMSDIT 360  
QY 256 SLVVOV---QNGRCALDELQ--TTOLLPTAICFRGRINQKVSNGENHVNMOVTNI- 308  
Db 361 DFVIRFPVQANRHDFNQETAGWSTPRPRITII--SENGSKLGTGVATDIYVGLP 418  
QY 309 NGTFPDPTGDVPAPLGTDFSGKLGVLRSORDH-----NACRSHDAVIATNSAKFTPKLG 364  
Db 419 DGMPTDTTICEELTPAG--DYS-----ITNGSGNDIATANAYDSADVIITNTFRGMYICG 471  
QY 365 AIQ 367  
Db 472 ALQ 474

RESULT 5  
US-10-670-695-2  
; Sequence 2, Application US/10670695  
; Publication No. US20040058316A1  
; GENERAL INFORMATION:  
; APPLICANT: Jensen, Wayne A.  
; APPLICANT: Lappin, Michael R.  
; APPLICANT: Rosen, David K.  
; APPLICANT: Andrews, Janet S.  
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE  
; FILE REFERENCE: DI-9-1  
; CURRENT APPLICATION NUMBER: US/10/670,695  
; CURRENT FILING DATE: 2003-09-25  
; PRIOR APPLICATION NUMBER: 09/521,738  
; PRIOR FILING DATE: 2000-03-09  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 671  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-10-670-695-2

Query Match 9.8%; Score 277.5; DB 12; Length 671;  
Best Local Similarity 31.4%; Pred. No. 4.1e-17;  
Matches 76; Conservative 38; Mismatches 97; Indels 31; Gaps 9;  
QY 30 EPVAGASIAAPLTGQNNIIDPW--IRLNFVQAPNGEFTVSPRNSPGEVLLNLELGPRLNPY 88  
Db 145 EPSAQMAATAADATGKSDSEWSEFSTSYNW----STSETQKILFKQSLGPLLPY 200  
QY 89 LAHLSRMVYAGGVEVQVLLAGNAFTAGKLVFAAVPPHFFPLENISPGQIT---MPPHVI 145  
Db 201 LEHLKLYVWAGSGSEVRSISGSGVFGKLAIAIVVPP-----GDPVQSTSMLOYPHVL 255  
QY 146 IDVRLPEVLLPLDVRNFFHYNQNEPRMELVAMLYTPLRSNGSGDDVFTVSC--RVL 203  
Db 256 FDARQVEPVIFCLPDLRSTLYHLMSDTD--TTSLVIMVYNDL--INPYANDANSSGCIVTVE 313  
QY 204 TRPSDFDFNVLVPPTLESKTKPFTLPILTIGELTNSRPP-----VPIDELYTSNPSLVV 259  
Db 314 TKPGDPDFKHLKPPG-----SMLTHGSIPLDKTSSSLWIGNRYMSDITDFVI 363  
QY 260 QP 261  
Db 364 RP 365

RESULT 6  
US-10-327-481A-38  
; Sequence 38, Application US/10327481A  
; Publication No. US20040001864A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Andrew M.Q.  
; APPLICANT: Burman, Alison J.  
; APPLICANT: Audonnet, Jean-Christophe F.  
; APPLICANT: Lombard, Michel F.A.  
; TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease  
; FILE REFERENCE: 454313-3178  
; CURRENT APPLICATION NUMBER: US/10/327,481A  
; CURRENT FILING DATE: 2002-12-20  
; PRIOR APPLICATION NUMBER: PCT/FR01/02042  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: FR 0008437  
; PRIOR FILING DATE: 2000-06-29  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: Patentin Ver. 2.1  
; SEQ ID NO 38  
; LENGTH: 1147  
; TYPE: PRT

QY 227 FFLPILITGELTNGRFPVPIDELYTSNESIAVQPQNGRCALDGLQGTTQLLPTAICSF 286

Dd 219 -TAPIMTQGSLYN-----DSLSTNDSKSLILLGSTPLDIAPGAVFQDLRLLSIDYSLG 270

Query Match	4.4%	Score 125.5;	DB 15;	Length 1147;
Best Local Similarity	22.1%;	Pred. No. 0.033;		
Matches	71;	Conservative	34;	Mismatches 139; Indels 77; Gaps 10;
QY	70	NSPGEVLNLELGP	LDNLYLAHL	SRMYNGVAGGVEQVLLAGNAFTAGKLVFAAAVPPHPF 129
DB	158	DKPFGYLKLEL	TFDHHGVFGLH	VDSEAYMRNGMDVEVSAGNQFNGGCLLVAMVPENKA 217
QY	130	LENISPGQITMP	PHVIVDRTLE	PVLLPLPDVRNRNFPHYNQNEPRMRLVAMLYTPLRSN 189
DB	218	FDTREKQITL	PHQFISPRIN	WTAHITVPYLVN--RYDQY----- 257
QY	190	GSDDVFTV	SCRVLT	TRSPDDFNVLPPTLESTKPTLPILTIGELTNSRFPPIDEL 249
DB	258	-----	-----	-----KKHKPTLVVMVSLVTSNTAADQIKV 285
QY	250	YTSNESLV	VQONGECAL	DGELQGTOLLPTAICSFRG-----RINQKSGENHVVN 302
DB	286	YAN-----	IAPTYVHVA--	GELPSKEGIFPVACADYGGLVTTDPKTADPVYVK--VTN 335
QY	303	MQVTNINGT	PFDPTGDV	PAPLGTDPF-----SGKLVGLVSQRDHONACRSHDAVIATNSAK 358
DB	336	PKNTYVGR--	FTNLLDVAE--	ACPFLRFDGDKPY--VVTRADDTLLAKFDSVLAACHMS 391
QY	359	FTPKLGAIG	IQGTWEED	DVHN 379
DB	392	NTVLSGIA	OYVTOYSG	TINLH 412

```

RESULT 7
US-09-867-932A-2
; Sequence 2, Application US/09867932A
; Publication No. US20030049825A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Han-You
; APPLICANT: Kuo, Tsun-Yung
; APPLICANT: Kuo, Hsiao-I
; APPLICANT: Yang, Huey-Lang
; TITLE OF INVENTION: NERVOUS NECROSIS VIRUS PROTEIN
; FILE REFERENCE: 08919-023001
; CURRENT APPLICATION NUMBER: US/09/867,932A
; CURRENT FILING DATE: 2001-05-30
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 338
; TYPE: PRT
; ORGANISM: Nervous necrosis virus
US-09-867-932A-2

```

Query Match	4.1%	Score 117;	DB 10;	Length 338;
Best Local Similarity	24.5%;	Pred. No. 0.031;		
Matches	80;	Conservative 49;	Mismatches 132;	Indels 66; Gaps 18;
Qy	8	AAPSDGAAGLV--EANNETWALEPVAGASIAAPL---	GQNNI-IDPWIRLNFVQAP	60
Db	10	AKPATTKAANPQRRRRNNRSNRTPDAPVSKASTVTCGRGTNDVHLSGMSRISQAVLP		69
Qy	61	NGBFTVSPRNSPGEVLLNLELCPELNPVLAHLSWYNGYAGGV---	EVQVLLAGNAFTAG	117
Db	70	AGTGT-----DGYVVVDATIVDDLLPRGHAAIRFORAYAVTELEFIQMPCAN--TGG		121
Qy	118	KLVFAAVPPHPFLENISP-GQITMEPHVII---DVTLEPVLPLPDVNNNEFFHYNQON		172
Db	122	GYVAGFLPD--PTDNDHTTGCALQATRGAVVAKWMSRTRV-----PQYTRILLWTSSCK		173
Qy	173	EPRM-----RLVAMLYTPLRSNGSGDDVFTVS--CRVLTRPSDDPYNLYPPTLESKTPK		226
Db	174	EORLTSPGRILLICV-----GNNTDVNVSVLCRWSVR-----SVPSLETPEET		218

```

227 FTLPILTIGELTNSRFVPIDELYTSPNBSLVVQPONGSCALDGEIQTQLLPTAICSF 286
219 -TAPLMTQGSLYN-----DSLSTNDSKSLILGSLPLDIAPDGAVFQDLRLLSIDYSLG 270
287 RGRINQKVSGENHWNMQVTNIN-GTP 312
271 TGDVDRAV-----YWHLKKFAGNAGTP 292

RESULT 8
US-10-282-122A-66335
; Sequence 66335, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIORITY APPLICATION NUMBER: 60/191,078
; PRIORITY FILING DATE: 2000-03-21
; PRIORITY APPLICATION NUMBER: 60/206,848
; PRIORITY FILING DATE: 2000-05-23
; PRIORITY APPLICATION NUMBER: 60/207,727
; PRIORITY FILING DATE: 2000-05-26
; PRIORITY APPLICATION NUMBER: 60/230,335
; PRIORITY FILING DATE: 2000-09-06
; PRIORITY APPLICATION NUMBER: 60/230,347
; PRIORITY FILING DATE: 2000-09-09
; PRIORITY APPLICATION NUMBER: 60/242,578
; PRIORITY FILING DATE: 2000-10-23
; PRIORITY APPLICATION NUMBER: 60/253,625
; PRIORITY FILING DATE: 2000-11-27
; PRIORITY APPLICATION NUMBER: 60/257,931
; PRIORITY FILING DATE: 2000-12-22
; PRIORITY APPLICATION NUMBER: 60/267,636
; PRIORITY FILING DATE: 2001-02-09
; PRIORITY APPLICATION NUMBER: 60/269,308
; PRIORITY FILING DATE: 2001-02-16
; Remaining prior application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66335
; LENGTH: 2468
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-10-282-122A-66335

Query Match 3.9%; Score 110.5; DB 12; Length 2468;
Best Local Similarity 18.7%; Pred. No. 3.1;
Matches 132; Conservative 64; Mismatches 246; Indels 265; Gaps 30;

QY 8 AAPSNDGACGLVP---EANNETWALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAPNGEF 64
DB 1234 AATTVDVAPPAPVIDPDSNGTTSTGTABGAKVI--LTDGNG-----NPITGET 1279
QY 65 TVSPRN-----SPCEVLNLELGPENLPYLAHLRSMYNGYAGGVEVQVLL---AGNAFTAG 117
DB 1280 TADGSGNWSFTPGTPLAN-----GTVVNAVAQDPAGNTPGQG 1316
QY 118 KLVEAAVPPHPHLENISPGQI---TMFPHVIDRTL-----EPVLLPLPVRNNFFHYN 169

```

Db 1317 STTDAVAPNTVVPNSGNLLNGTAEPGSTV---TLTDGNGNPICQTTADSGNWSFTP 1373  
Qy 170 QONEPRMLVAMLYTPLRNSGDDVFTVSCRVLTRPSPDFDFNYLVPPTLSESKTKPFTL 229  
Db 1374 GSQLPNGTVVNTASDAAGNTSLPATTVDSSLSPIPOVDPSNGVISCTADAGN---TI 1430  
Qy 230 PILT-----IGELT-----NSRF-----PVP-----IDELYTSPNE----- 255  
Db 1431 -IITDNGNPIQOVTDAGSNWSFTPGIPLPDGTVVNVVARSFNSVDSAPAVITVDGVAP 1489  
Qy 256 -SLVOPQNGRCALDELQGTITQLLPTAICSFRG--RINOKVSGENHVNWM-----QV 305  
Db 1490 AAPVIDPSNGT-----EISGTAEAGATVILTDGGNPIQATADSGNMTFTPSTPLANG 1544  
Qy 306 TNINGTFFDPTGDV-----PAPL----- 323  
Db 1545 TVINAVAOQDPAGNTSGPASVTDATAPAPVINPNSGVVISGTAEAGATVILTDGNGNFI 1604  
Qy 324 -----GTPDFSGKLFVLSQRDHDNACRSHDAVI-----ATNSAKFTPKLGA 365  
Db 1605 GOVTADSGNWSFTPGTPLANGSVINALAQDAAGNNSPTSATVDSLAPAAPVIDPSNGS 1664  
Qy 366 IOIGTWEEDDVHINQPTKFTPVGLFENEGFNOWTLP-----NYSG 405  
Db 1665 VIAGTAEAGATVILTDGNGNPIQOVTDAGSNWSFTPGTPLSGTNGTVVNAVAQDAAGNTSG 1724  
Qy 406 ALTNMGLAPPVAPTF-----PGSQILF-----FRSH 432  
Db 1725 PVSTTVDAVAPATPVIDPSNGVELSGTAEPGVRVILTDGNGNPIQOTLADSGNWSFTPG 1784  
Qy 433 IPLKGGVA-----DPVIDCLLPQEWIQLHYESAPSQSDVALIR-----FTNPDTGRVLP 482  
Db 1785 TPLANGTVVNAVAQDPA-----GNTSGPASTTVDTVAPATPVINPNSGVI- 1830  
Qy 483 EAKLHRSYITVANTGSRPIVVPANGYFRFDTVWNOFYSLAPMGTCN 529  
Db 1831 -----TGTAEVGAKVILTDGNG-----NPIGETTADGSGN 1860

RESULT 9  
US-10-246-330-4  
; Sequence 4, Application US/10246330  
; Publication No. US20030166030A1  
; GENERAL INFORMATION:  
; APPLICANT: O'Toole, George A.  
; TITLE OF INVENTION: METHODS TO STUDY AND MECHANISMS OF  
; FILE REFERENCE: 14537-002001  
; CURRENT APPLICATION NUMBER: US/10/246,330  
; PRIOR FILING DATE: 2002-09-18  
; PRIOR APPLICATION NUMBER: US 60/323,241  
; NUMBER OF SEQ ID NOS: 32  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 4  
; LENGTH: 2468  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-10-246-330-4

Query Match 3.9%; Score 110.5; DB 14; Length 2468;  
Best Local Similarity 18.7%; Pred. No. 3.1;  
Matches 132; Conservative 64; Mismatches 246; Indels 265; Gaps 30;

Qy 8 AAPNDGAAGLVP-----EANNETWALEPVGASIAAPLTQCNIIIDPWIRLNFVQAPNGEF 64  
Db 1234 AATTVDVAPAPVIDPSNGTITSGTAEAGAKVI--LTDGNG-----NPIGET 1279  
Qy 65 TVSPRN-----SPCEVLLNLELGPENLPYLAHLSRMYNGVAGVEVQVLL---AGNAFTAG 117  
Db 1280 TADSGNWSFTPGTPLAN-----GTVVNAVAQDPAGTNGPQG 1316

Qy 118 KLVFAAAPHPPPLENISPGQI---TMPPHVIDVTL-----EPVLLPLPDVNRNPPHYN 169  
Db 1317 STTDAVAPNTVVPNSGNLLNGTAEPGSTV---TLTDGNGNPICQTTADSGNWSFTP 1373  
Qy 170 QONEPRMLVAMLYTPLRNSGDDVFTVSCRVLTRPSPDFDFNYLVPPTLSESKTKPFTL 229  
Db 1374 GSQLPNGTVVNTASDAAGNTSLPATTVDSSLSPIPOVDPSNGVISCTADAGN---TI 1430  
Qy 230 PILT-----IGELT-----NSRF-----PVP-----IDELYTSPNE----- 255  
Db 1431 -IITDNGNPIQOVTDAGSNWSFTPGIPLPDGTVVNVVARSFNSVDSAPAVITVDGVAP 1489  
Qy 256 -SLVOPQNGRCALDELQGTITQLLPTAICSFRG--RINOKVSGENHVNWM-----QV 305  
Db 1490 AAPVIDPSNGT-----EISGTAEAGATVILTDGGNPIQATADSGNMTFTPSTPLANG 1544  
Qy 306 TNINGTFFDPTGDV-----PAPL----- 323  
Db 1545 TVINAVAOQDPAGNTSGPASVTDATAPAPVINPNSGVVISGTAEAGATVILTDGNGNFI 1604  
Qy 324 -----GTPDFSGKLFVLSQRDHDNACRSHDAVI-----ATNSAKFTPKLGA 365  
Db 1605 GOVTADSGNWSFTPGTPLANGSVINALAQDAAGNNSPTSATVDSLAPAAPVIDPSNGS 1664  
Qy 366 IOIGTWEEDDVHINQPTKFTPVGLFENEGFNOWTLP-----NYSG 405  
Db 1665 VIAGTAEAGATVILTDGNGNPIQOVTDAGSNWSFTPGTPLSGTNGTVVNAVAQDAAGNTSG 1724  
Qy 406 ALTNMGLAPPVAPTF-----PGSQILF-----FRSH 432  
Db 1725 PVSTTVDAVAPATPVIDPSNGVELSGTAEPGVRVILTDGNGNPIQOTLADSGNWSFTPG 1784  
Qy 433 IPLKGGVA-----DPVIDCLLPQEWIQLHYESAPSQSDVALIR-----FTNPDTGRVLP 482  
Db 1785 TPLANGTVVNAVAQDPA-----GNTSGPASTTVDTVAPATPVINPNSGVI- 1830  
Qy 483 EAKLHRSYITVANTGSRPIVVPANGYFRFDTVWNOFYSLAPMGTCN 529  
Db 1831 -----TGTAEVGAKVILTDGNG-----NPIGETTADGSGN 1860

RESULT 10  
US-09-349-058-38  
; Sequence 38, Application US/09349058  
; Publication No. US20030056229A1  
; GENERAL INFORMATION:  
; APPLICANT: Crabtree, Gerald R.  
; APPLICANT: No. US20030056229A1thiop, Jeffrey P.  
; TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES  
; FILE REFERENCE: APV-332.12  
; CURRENT APPLICATION NUMBER: US/09/349,058  
; CURRENT FILING DATE: 1998-07-07  
; EARLIER APPLICATION NUMBER: 08/260,174  
; EARLIER FILING DATE: 1994-06-13  
; NUMBER OF SEQ ID NOS: 52  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 38  
; LENGTH: 716  
; TYPE: PRT  
; ORGANISM: human  
US-09-349-058-38

Query Match 3.9%; Score 110; DB 10; Length 716;  
Best Local Similarity 21.5%; Pred. No. 0.48;  
Matches 109; Conservative 62; Mismatches 201; Indels 134; Gaps 26;

Qy 111 GNAFTAGKLVFAAAPPVPH---FPLENISPGQITMPPHVI-----DVTLEPVLPLPDVNRN 163  
Db 28 GPAPRAGGTMTKSABEEHYGASSNVSPALPLPTAHTSLPAPCHNLQSTSTFGIIPADHPS 87  
Qy 164 NF-----FHYNQ-----NEPRMLVAMLYTPLRNSGDDVFTVSCRVL 203

Db 88 GYGAALDCCPAGYFLSSGHTPDGAPALSPRIITSCILGYHNNQFFHDVEVEDVLP 147  
Qy 204 TRPSDFDNFVLPPTLESKTKPFTL-PILTIGELT-----NSRFPVIDELYTSPN 254  
Db 148 SKRSP--STATLSLSLEAYRDPSCLSPASSLSRSCNSEASSYESYSPYASPTSPW 205  
Qy 255 ESLVQOPN-----GRCALDGLQGTQTLPTAICSFGRINQKVSNGENHVMNM 303  
Db 206 QSPCVSPKTTDPEEGFPRGLGACTLLGSPQSPSTSPRASVTEESWLGARSRPASPCNK 265  
Qy 304 QVTNING--TPDTP-TGDVPAPLGTDFDS---GKLFGLVLSQRDHDNACRSHDAVIATNSA 357  
Db 266 RYKSLNGRQPPSPHSPPTSPHSGPRSVTDDSLGNTTQ-----YTSSAIVAAINA 318  
Qy 358 KFTPKLGAIOIGTWEEDDVIH-----NQP-----TKFTPVCLFENEGFNQWTLPNYSCA 406  
Db 319 LTTD--SSLDLG---DGVPKSRKTTLEQPPSVALKVEPVG-----354  
Qy 407 LTLNMGGLAPPVAPTFPEGEQILFRSHIPLKGGVADPVIDCLLPQEWIOHLYOESAPS---463  
Db 355 --EDLGSPPPPADAFEDYSSP--QHI-RKGGFCQDYL--AVP-----QHPYQWAKPKPLS 403  
Qy 464 -----QSDVALIRFTNPTGRVLFPEAKL-HRSGYIT-----VANTGSRPIVVP 505  
Db 404 PTSYMSPTLPALDQWLPSSHGPGYELRIEVQPKSHHRAHYETEGSRGAVKASAGGHP1-VQ 462  
Qy 506 ANGYFREDTWNQFYSLAPMGTCNGR 531  
Db 463 LHGYLENEPLMLQLF-----IGTADDR 484

RESULT 11  
US-10-040-430-38  
; Sequence 38, Application US/10040430  
; Publication No. US20030049641A1  
; GENERAL INFORMATION:  
; APPLICANT: Crabtree, Gerald R.  
; APPLICANT: No. US20030049641A1throp, Jeffrey P.  
; APPLICANT: Ho, Steffan M.  
; APPLICANT: Flanagan, William M.  
; TITLE OF INVENTION: NF-AT POLYPEPTIDES AND POLYNUCLEOTIDES AND SCREENING  
; FILE REFERENCE: APV-008.04  
; CURRENT APPLICATION NUMBER: US/10/040.430  
; CURRENT FILING DATE: 2002-01-07  
; PRIOR APPLICATION NUMBER: US/09/232.346  
; PRIOR FILING DATE: 1999-01-15  
; PRIOR APPLICATION NUMBER: 08/507,032  
; PRIOR FILING DATE: 1995-07-31  
; PRIOR APPLICATION NUMBER: 08/228,944  
; PRIOR FILING DATE: 1994-04-18  
; PRIOR APPLICATION NUMBER: 07/749,385  
; PRIOR FILING DATE: 1991-08-22  
; PRIOR APPLICATION NUMBER: 08/260,174  
; PRIOR FILING DATE: 1994-06-13  
; PRIOR APPLICATION NUMBER: 08/124,981  
; PRIOR FILING DATE: 1993-09-20  
; NUMBER OF SEQ ID NOS: 62  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 38  
; LENGTH: 716  
; TYPE: PRT  
; ORGANISM: human  
US-10-040-430-38  
Query Match 3.9%; Score 110; DB 14; Length 716;  
Best Local Similarity 21.5%; Pred. No. 0.48;  
Matches 109; Conservative 62; Mismatches 201; Indels 134; Gaps 26;  
Qy 111 GNAFTAGKLVFAAVPPH--FPLENISPGQITMPPHVI-----DVRTLEPVLPLPDVNR 163  
Db 28 GPAPRAGGTMKSAEEHYGASSNVSPALPLPTAHSTLPACNHLQSTPGIIPADHPS 87

Qy 164 NF-----FHYNQ-----NEPRMLVAMLYTPLRNSGSGDDVFTVSCRVL 203  
Db 88 GYGAALDCCPAGYFLSSGHTPDGAPALSPRIITSCILGYHNNQFFHDVEVEDVLP 147  
Qy 204 TRPSDFDNFVLPPTLESKTKPFTL-PILTIGELT-----NSRFPVIDELYTSPN 254  
Db 148 SKRSP--STATLSLSLEAYRDPSCLSPASSLSRSCNSEASSYESYSPYASPTSPW 205  
Qy 255 ESLVQOPN-----GRCALDGLQGTQTLPTAICSFGRINQKVSNGENHVMNM 303  
Db 206 QSPCVSPKTTDPEEGFPRGLGACTLLGSPQSPSTSPRASVTEESWLGARSRPASPCNK 265  
Qy 304 QVTNING--TPDTP-TGDVPAPLGTDFDS---GKLFGLVLSQRDHDNACRSHDAVIATNSA 357  
Db 266 RYKSLNGRQPPSPHSPPTSPHSGPRSVTDDSLGNTTQ-----YTSSAIVAAINA 318  
Qy 358 KFTPKLGAIOIGTWEEDDVIH-----NQP-----TKFTPVCLFENEGFNQWTLPNYSCA 406  
Db 319 LTTD--SSLDLG---DGVPKSRKTTLEQPPSVALKVEPVG-----354  
Qy 407 LTLNMGGLAPPVAPTFPEGEQILFRSHIPLKGGVADPVIDCLLPQEWIOHLYOESAPS---463  
Db 355 --EDLGSPPPPADAFEDYSSP--QHI-RKGGFCQDYL--AVP-----QHPYQWAKPKPLS 403  
Qy 464 -----QSDVALIRFTNPTGRVLFPEAKL-HRSGYIT-----VANTGSRPIVVP 505  
Db 404 PTSYMSPTLPALDQWLPSSHGPGYELRIEVQPKSHHRAHYETEGSRGAVKASAGGHP1-VQ 462  
Qy 506 ANGYFREDTWNQFYSLAPMGTCNGR 531  
Db 463 LHGYLENEPLMLQLF-----IGTADDR 484

RESULT 12  
US-10-116-275-182  
; Sequence 182, Application US/10116275  
; Publication No. US20030211476A1  
; GENERAL INFORMATION:  
; APPLICANT: Elan Pharmaceutical Technology  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Brayden, David  
; APPLICANT: Byrne, Daragh  
; APPLICANT: Lamkin, Imelda  
; APPLICANT: Higgins, Lisa  
; TITLE OF INVENTION: Genetic Analysis of Peyer's Patches and M Cells and Methods and  
; TITLE OF INVENTION: Compositions targeting Peyer's Patches and M Cell Receptors  
; FILE REFERENCE: E1067/20087  
; CURRENT APPLICATION NUMBER: US/10/116.275  
; CURRENT FILING DATE: 2002-10-04  
; NUMBER OF SEQ ID NOS: 349  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 182  
; LENGTH: 716  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-116-275-182  
Query Match 3.9%; Score 110; DB 15; Length 716;  
Best Local Similarity 21.5%; Pred. No. 0.48;  
Matches 109; Conservative 62; Mismatches 201; Indels 134; Gaps 26;  
Qy 111 GNAFTAGKLVFAAVPPH--FPLENISPGQITMPPHVI-----DVRTLEPVLPLPDVNR 163  
Db 28 GPAPRAGGTMKSAEEHYGASSNVSPALPLPTAHSTLPACNHLQSTPGIIPADHPS 87  
Qy 164 NF-----FHYNQ-----NEPRMLVAMLYTPLRNSGSGDDVFTVSCRVL 203  
Db 88 GYGAALDCCPAGYFLSSGHTPDGAPALSPRIITSCILGYHNNQFFHDVEVEDVLP 147  
Qy 204 TRPSDFDNFVLPPTLESKTKPFTL-PILTIGELT-----NSRFPVIDELYTSPN 254  
Db 148 SKRSP--STATLSLSLEAYRDPSCLSPASSLSRSCNSEASSYESYSPYASPTSPW 205





Db 609 V-TVYASIR-----QASPIRLASVTALIESVNGKTVTLELDNGAG 649  
Qy 297 ENHVNMQVTNINGTFFDPTGVPAPLGTDFSGKLF---GVLSQRDHDNACRSHDAVIA 353  
Db 650 ADATKNDGVYSRFFTAFDANG-----RYSVKIWAALGGVTSRDRQ- 688  
Qy 354 TNSAKFTPKLGAIQIGTWEED-DVHINOPTKFTPVGLFENEGFNOWTLIPNYSGA--LTLN 410  
Db 689 ----AAPPKNRMYIDGMIEDGEVRNMPRETSTVQDKQLCFRTS-----SGGSFVATN 740  
Qy 411 MGLAPVAPTFPGEQILFFRSHIPKGGVADVIDCLLPQEWIOHLYOESAPSDVALI 470  
Db 741 VPAAPIPDLFPCCQITDLKASIQGN-----LVNLTW-----TAPGD----- 778  
Qy 471 RFTNPTGRV-----LFEAKLHRSYITVANTGSRPIVVPANGYERFDTWVNOFY 520  
Db 779 ---DYDHGRASNYIIRMSTSIVDLRDHFNTSLQVNTTGLIPKEASSEIFEELGGNTF- 834  
Qy 521 SLAPMGTTGNG 530  
Db 835 -----GNG 837

Qy 246 IDELYTSPNESLVOPQNGRCALDGELOQTOLLPTAICSPRGRINOKV-----SG 296  
Db 609 V-TVYASIR-----QASPIRLASVTALIESVNGKTVTLELDNGAG 649  
Qy 297 ENHVNMQVTNINGTFFDPTGVPAPLGTDFSGKLF---GVLSQRDHDNACRSHDAVIA 353  
Db 650 ADATKNDGVYSRFFTAFDANG-----RYSVKIWAALGGVTSRDRQ- 688  
Qy 354 TNSAKFTPKLGAIQIGTWEED-DVHINOPTKFTPVGLFENEGFNOWTLIPNYSGA--LTLN 410  
Db 689 ----AAPPKNRMYIDGMIEDGEVRNMPRETSTVQDKQLCFRTS-----SGGSFVATN 740  
Qy 411 MGLAPVAPTFPGEQILFFRSHIPKGGVADVIDCLLPQEWIOHLYOESAPSDVALI 470  
Db 741 VPAAPIPDLFPCCQITDLKASIQGN-----LVNLTW-----TAPGD----- 778  
Qy 471 RFTNPTGRV-----LFEAKLHRSYITVANTGSRPIVVPANGYERFDTWVNOFY 520  
Db 779 ---DYDHGRASNYIIRMSTSIVDLRDHFNTSLQVNTTGLIPKEASSEIFEELGGNTF- 834  
Qy 521 SLAPMGTTGNG 530  
Db 835 -----GNG 837

Search completed: June 1, 2004, 14:04:42  
Job time : 36.2005 secs

RESULT 15  
US-10-369-214-132  
; Sequence 132, Application US/10369214  
; Publication No. US20030232037A1  
; GENERAL INFORMATION:  
; APPLICANT: Groot, Pieter C.  
; APPLICANT: Berghenhegouwen van, Bram J.  
; APPLICANT: Oosterhout van, Antoon J.M.  
; TITLE OF INVENTION: Genes involved in immune related responses observed  
; TITLE OF INVENTION: with asthma  
; FILE REFERENCE: P53837US00  
; CURRENT APPLICATION NUMBER: US/10/369,214  
; CURRENT FILING DATE: 2003-02-15  
; PRIOR APPLICATION NUMBER: EP 00202867.8  
; PRIOR FILING DATE: 2000-08-16  
; PRIOR APPLICATION NUMBER: PCT/NL01/00610  
; PRIOR FILING DATE: 2001-08-16  
; NUMBER OF SEQ ID NOS: 139  
; SOFTWARE: PatentIn ver. 2.1  
; SEQ ID NO 132  
; LENGTH: 913  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (1)..(913)  
; OTHER INFORMATION: /note="Calcium-activated chloride channels Gob-5"  
US-10-369-214-132

Query Match 3.8%; Score 109; DB 15; Length 913;  
Best Local Similarity 21.5%; Pred. No. 0.87;  
Matches 118; Conservative 60; Mismatches 208; Indels 164; Gaps 25;  
Qy 31 PVAGASIAAPLTGQNNIIDPWIRLNFVQAPNGEFTVSPRSPGVEVLNLELGPENPYLA 90  
Db 402 PTDGSEIVLLTDGEDNTISSCFDL-----VKQSGAIIHTVALGPAAKELE 447  
Qy 91 HLSRMVYAGVGEVQVLLAG--NAFTAGKLVFAAVPPH-FPLE-----NISPGQITWFP 143  
Db 448 QLSKMTGGLQTVSSDOVQNGNLVDAPALSSGNAIAQHSIQLESRGVNLQNNQ-WWNGS 506  
Qy 144 VIID-----VRTLEPVLLPLPD---VRNFFHYNOQNEPRMLVAMLYTPLRSN 189  
Db 507 VIVDSSVGKDTLFLITWTHPTIFIDWPSGVEQNGFILDITTK-----VAYLQVPGTAK 561  
Qy 190 -GSGDDVFTVSCRVLTRSPDFDNVLPPTLESKTKPFTLPILTIGELTN---SRFPVP 245  
Db 562 VGFWKYSIQASSQTTLT-----TVTSRAASATLPIIVTPVKNKTKKFPSP 608

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.46309 Seconds  
(without alignments)  
4310.252 Million cell updates/sec

Title: US-09-926-799-8  
Perfect score: 2843  
Sequence: 1 MKMASNDAPSDGAGLVLP.....VNQFVSLAPMGTGNGRRRVQ 535

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1925	67.7	539	1 COAT LORDV	P54635 lordsdale v
2	1174	41.3	546	1 COAT SOUV3	Q04542 southampton
3	325.5	11.4	2344	1 POLN_RHDV	P27410 rabbit hemo
4	289	10.2	702	1 COAT SMSV1	P36284 san miguel
5	284	10.0	668	1 COAT FCVF4	P27405 feline cali
6	277.5	9.8	671	1 COAT FCVF9	P27406 feline cali
7	273.5	9.6	703	1 COAT SMSV4	P36285 san miguel
8	272.5	9.6	668	1 COAT FCVF6	P27404 feline cali
9	262	9.2	2208	1 POLN_MANCV	Q69014 manchester
10	142	5.0	2193	1 POLG_CX16T	Q9qf31 c genome po
11	137	4.8	2193	1 POLG_CX16G	Q65900 c genome po
12	130	4.6	2206	1 POLG_POL32	P06209 poliovirus
13	128	4.5	2194	1 POLG_HE701	P32537 h genome po
14	125.5	4.4	2333	1 POLG_FMDV1	P03306 f genome po
15	125.5	4.4	4705	1 FAT2_DROME	Q9vw71 drosophila
16	122	4.3	980	1 CLV1_ARATH	Q98yq8 arabidopsis
17	121.5	4.3	2332	1 POLG_FMDV0	P03305 f genome po
18	120.5	4.2	2205	1 POLG_POL2W	P23069 p genome po
19	120	4.2	1011	1 POLG_FMDVT	P15072 foot-and-mo
20	119.5	4.2	2336	1 POLG_FMDVZ	P49303 f genome po
21	118	4.2	2193	1 POLG_HE71M	P06207 h genome po
22	115	4.0	2332	1 POLG_FMDVA	P03308 f genome po
23	114	4.0	2207	1 POLG_POL2L	P06210 p genome po
24	112	3.9	2206	1 POLG_POL3L	P03302 poliovirus
25	109	3.8	2193	1 POLG_EC09H	P06849 e genome po
26	109	3.8	2201	1 POLG_CXA9	P21404 c genome po
27	108.5	3.8	808	1 POLG_HP4V3	Q02381 hepatitis a
28	108	3.8	2193	1 POLG_HE71B	Q66478 h genome po
29	107	3.8	788	1 BCSB_XANAC	P58933 xanthomonas
30	106.5	3.7	1739	1 DOTL_HUMAN	O8tek3 homo sapien
31	106	3.7	943	1 NFCL_HUMAN	O95644 homo sapien
32	104	3.7	2184	1 POLG_EC01F	Q91734 e genome po
33	102.5	3.6	1698	1 41_DROME	Q9v8r9 drosophila

34 101.5 3.6 852 1 POLG\_HP4VC P06442 hepatitis a  
35 101.5 3.6 2185 1 POLG\_SVDVU P13900 s genome po  
36 101.5 3.6 2196 1 POLG\_EC05N Q9ylj1 e genome po  
37 101.5 3.6 2206 1 POLG\_POL1M P03299 p genome po  
38 101.5 3.6 2230 1 POLG\_HP4VS P14553 sinian hepa  
39 100.5 3.5 604 1 RCOL\_NEUCR P78706 neurospora  
40 100.5 3.5 955 1 AGAB\_VIBS7 P48840 vibrio sp.  
41 100.5 3.5 2208 1 POLH\_POL1M P03300 p genome po  
42 100.5 3.5 2226 1 POLG\_HP4V2 P26580 hepatitis a  
43 100.5 3.5 2226 1 POLG\_HP4V4 P26581 hepatitis a  
44 100.5 3.5 2227 1 POLG\_HP4VH P06617 hepatitis a  
45 100.5 3.5 2227 1 POLG\_HP4VL P06441 hepatitis a

## ALIGNMENTS

## RESULT 1

ID COAT LORDV STANDARD; PRT; 539 AA.  
AC P54635;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE Coat protein (Capsid protein).  
OS Lordsdale virus (Human enteric calicivirus).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=82658;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96005060; PubMed=7561776;  
RA Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;  
RT "Human enteric Caliciviridae: the complete genome sequence and  
RT expression of virus-like particles from a genetic group II small  
RT round structured virus".  
RL J. Gen. Virol. 76:2349-2355(1995).  
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; X86557; CAA60255.1; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
KW Coat protein; Glycoprotein.  
FT CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 539 AA; 58775 MW; 585C63E7F2C5FD21 CRC64;

Query Match 67.7%; Score 1925; DB 1; Length 539;

Best Local Similarity 65.7%; Pred. No. 6.3e-137;

Matches 355; Conservative 79; Mismatches 96; Indels 10; Gaps 5;

QY 1 MKMASNDAPSDGAGLVLPPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60

Db 1 MKMASNDANPSDGSANLVPEVNEVMALEPVGNAIAAPVAGQNVIDPWIRNFVQAP 60

QY 61 NGEFTVSPRNSPGEVLLNLELGPBLNLYLAHLSRMYNGYAGVEVQVLLAGNAFTAGKLV 120

Db 61 GGEFTVSPRNPAGBEILWSAPLGPDNLNLYLAHLSRMYNGYAGVEVQVLLAGNAFTAGKVI 120

QY 121 FAAYPPHPLENIQSGQITMEPHVITDVRTLEPVLLPLPDVNRNPFHYNQNEPRMRLVA 180

Db 121 FAAYPPNPFTGLSPSQVTMPFHIIVDRQLEPVLIPLDVNRNPFHYNQANDSTKLIA 180

1 MMASKADPQADGASGAGQLVPEVNTADPLPMEPVAGPTTAVATAGQVNMIDPWIWNF 60

57 VOAPNGEFTVSPRNSPEVLLNLELGPALNLAHLSRMVNGVAGGVEVVOVLGNAFTA 116

61 VQSPQCEFTISPNNTPGDILFDLQGLPPLNPLFSLHLSQMTNGVGNMRVILLAGNAFSA 120

117 GKLVFAAPPHFPLENISPGQITMFPFVHVIDVRLTEPVLPLPDVRRNPFHYNOQNEPRM 176

121 GKIIIVCCVPPGETSSSLIAQATLFPVHIAVDRTELEPIEMPLEDVRNVLYHTN-DNQPTM 179

177 RLVAMLYTLPLRS-NGSGD-DVFTVSCRVLTPSPDFDFFNYLVPTLESKTKPFTLPILTI 234

180 RLVCMLYTLPLRTGCGSGNSDSFVWAGRVLTAPSSDFSLFLVPTTIEQKTRAFVFNPL 239

235 GELTNSRFPVPTDELYTSPNESLVVQPNQGCALDGELOQTTLPTAICSFRGRINOKV 294

240 QTLNSRFPVPTDELYTSPNESLVVQPNQGCALDGELOQTTLPTAICSFRGRINOKV 299

295 SGENHVMNQVNTINGPTDPTGDPAPLPGTDPFSGKLFVLSQRDHONACRSHDAV--- 351

300 R-----TLNLTEDVCKPF-MAFDSPAPVGFDF-GKCDWHMRISKTNNNTGSGDPMRSV 351

352 -IATNAKFTPKLGAIQ-----IGTWEDDDVHINOPTKFTPTVGLFENEFGNQW 398

352 SVQTNVQGFVPHLGSIQDFEVNHNPTGDIYIGIE-----WISQPS--TPPG---TDINLW 401

399 TLPNYSGLATLNMGLAPPVAPTPPGQILFFRSHIPLKGGVADP-VIDCLLPQEWIHL 457

402 EIPDYGSSLSQAANLAPPVFPFGFGEALVVFSAFEGPNNRSPANDVPCLLPOEYITHFV 461

458 QESAPQSODVALLRFTNPTDGRVLFKALHRSYITVA--NTGSRPIVVPVANGYFRFTW 515

462 SEQAPTMGDAALLHYDDPTNRNLGFBKLYPGGILYTCVNGVAGPQQLPLNGVFLVSW 521

516 VNQFYSIAPMGTCN-----GRRRV 534

522 VSRFYQLKXPGVTASTARGRLGVRI 546

RESULT 3

POLN RHDV STANDARD; PRT: 2344 AA.

ID POLN RHDV STANDARD; PRT: 2344 AA.

AC P27410; (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Non-structural polyprotein [Contains: RNA-directed RNA polymerase (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like protein); Coat protein].

DE Rabbit hemorrhagic disease virus (RHDV).

OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae; Lagovirus.

OC NCBI\_TaxID=11976;

OX NCBI\_TaxID=11976;

RN [1] SEQUENCE FROM N.A.

RP MEDLINE=91361557; PubMed=1840711;

RA Meyers G, Wirblich C, Thiel H.-J.;

RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide sequencing of a calicivirus genome.";

RL Virology 184:664-676(1991).

CC -I- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.

CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA)(N).

CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

181 MLYTPLRNGSGDDVFTVSCRVLTSPSPDFDFFNYLVPTLESKTKPFTLPILTI 240

181 MLYTPLRANNAGDDVFTVSCRVLTSPSPDFDFFNYLVPTLESKTKPFTLPILTI 240

241 RFPVVIDELYTSPNESLVVQPNQGCALDGELOQTTLPTAICSFRGRINOKVSGNHV 300

241 RFPVVIDELYTSPNESLVVQPNQGCALDGELOQTTLPTAICSFRGRINOKVSGNHV 300

301 WNNQVNTINGPTDPTGDPAPLPGTDPFSGKLFVLSQRDH-DNACRSHDAVIATNSAKF 359

299 YTMNLASQNSYNDPTEIPAPLPGTDPFSGKLFVLSQRDH-DNACRSHDAVIATNSAKF 358

360 TPKLGAIQGTWEEDVHINOPTKFTPKVGLF-----ENEGFNQWTLNYSGLATLNMGL 413

359 TPKLGSVQFTTNDNDFQAGQNTKFTPVGVIOGDHQNKP-QQWLSLNYSGRTGHNVL 417

414 APPVATFPGEQILFFRSHIPLKGGVADPVIDCLLPQEWIHLYESAPSDVALLRFT 473

418 APAVATFPGEQILFFRSHIPLKGGVADPVIDCLLPQEWIHLYESAPSDVALLRFT 477

474 NPDTRGVLFKALHRSYITVA NTGSRPIVVPVANGYFRFTWVNQFYSIAPMGTCNGRRR 533

478 NPDTRGVLFKALHRSYITVA NTGSRPIVVPVANGYFRFTWVNQFYSIAPMGTCNGRRR 537

RESULT 2

ID COAT SOUV3 STANDARD; PRT: 546 AA.

AC Q04542; (Rel. 30, Created)

DT 01-OCT-1994 (Rel. 30, Last sequence update)

DT 01-OCT-1994 (Rel. 30, Last annotation update)

DE Coat protein (Capsid protein).

OS Southampton virus (serotype 3).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae; Norovirus.

OC NCBI\_TaxID=37129;

RN [1] SEQUENCE FROM N.A.

RP MEDLINE=93142023; PubMed=8380940;

RA Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;

RT "Sequence and genome organization of a human small round-structured (Norwalk-like) virus.";

RL Science 259:516-519(1993).

CC -I- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: L07418; AAA92984.1; --

DR FIR; B37491; B37491.

DR InterPro; IPR004005; Calici\_coat.

DR InterPro; IPR008975; Viral\_cap\_coat.

DR Pfam; PF00915; Calici\_coat; 1.

KW Coat protein; Glycoprotein.

FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).

SQ SEQUENCE 546 AA; 58774 MW; C82BZA85AD4B05EA CRC64;

Query Match 41.3%; Score 1174; DB 1; Length 546;

Best Local Similarity 46.2%; Pred. No. 1.5e-80;

Matches 261; Conservative 82; Mismatches 172; Indels 50; Gaps 16;

1 MMASNDAAASNDGAAG--LVPEANN-ETMALEPVAGASIAPLTGGNNIDPWLRLNF 56

```

CC EMBL; M67473; AAA47285.1; -
DR PIR; A41039; RRRWRH.
DR PDB; 1KHV; 16-JAN-02.
DR PDB; 1KHW; 16-JAN-02.
DR MEROPS; C24.001; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR000317; Peptidase_C24.
DR InterPro; IPR000605; RNA_helicase.
DR InterPro; IPR000709; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00916; 2CENDOPTASE.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.
FT CHAIN ? 508 7707 HELICASE (P2C-LIKE).
FT CHAIN 1109 1251 PROTEASE P3C.
FT CHAIN 71497 71625 RNA-DIRECTED RNA POLYMERASE.
FT CHAIN ? 2344 COAT PROTEIN.
FT NP BIND 522 529 ATP (POTENTIAL).
FT ACT_SITE 1212 1212 PROTEASE (POTENTIAL).
FT ACT_SITE 1227 1227 PROTEASE (POTENTIAL).
SQ SEQUENCE 2344 AA; 257066 MW; 1454C24F81E9212 CRC64;

Query Match 11.4%; Score 325.5; DB 1; Length 2344;
Best Local Similarity 26.3%; Pred. No. 5.2e-16;
Matches 152; Conservative 73; Mismatches 227; Indels 125; Gaps 30;

QY 8 AAPNDGAAGLVPEAN---NETMALEP-----VAGASIAAPLTGQNNI 47
DB 1772 AAPQGE-AAGTATTASVPGTTGMDPGVVATTSVITAENSASITAGIGGP-PQOVQ 1829

QY 48 IDPWRLNFVQAPNGEFTVSPNSQGVLLNLELGPENLYLAHLRMYNGVAGGVEQV 107
DB 1830 QETW-RTNFY--NDVFTWSVADAPGSIITYVQSPNNPFTAVLSQMYWAGGQFRF 1886

QY 108 LLAGNAFTAGKLVFAAVPPHPLENISPG-QITMPPHVIDVRLPVLPLPDVRNPF 166
DB 1887 IVAGSGVFGRLVRAVIPGI---EIGPGLVRFQPHVVIDARSLPVTITPDLRPNY 1943

QY 167 HYNQNEPRM--RLVAMLYTPLRNSGDDVFTVSCRVLTRPSPDPFNLYPPTLESKT 224
DB 1944 H--PTGDPGLVPTLVLSVNNL-INPFGGSTSAIQVTVETRPSEDPEFVMIRAPS--SKT 1998

QY 225 ----KP---FTPLTIGLTNSRFPVPIDELTYSNESLV-VQPNG---RCALDGELO 273
DB 1999 VDSISPAGLTTPFLVLT-GVGNDRW-----NGQIVGLQPPGFGFSTCNRHWNLN 2046

QY 274 GTTQLLPTA----ICSFGRINQKVSNGENHVMNQVTNINGTFFDPTGVDPAPLPTPDFS 329
DB 2047 GSTYGHSSPRFGDIDHRRSGASVSGSNATNVLQFTYANAGSAIDNPISQV-APDGPDM 2105

QY 330 GKLFGLVLSQRDHDNACRSHDAVIATNSAKFTPKLGAIQIGT-WEEDDVHINOPT-KFTPV 387
DB 2106 FVFP-----NGPGIPAAQWVGFGAIWNSGAPNVTVOAYEL 2143

QY 388 GLFENEGNQWLTLPVNSGALTLMNGLAPVAPFTFQGEQILPFRSHIPLKGGVADPIDCL 447
DB 2144 GFATGAPGNLQPTNTTSGATVAKSIYAVVTGTAQNPAGLF-----VMASGIISTPNASAI 2199

QY 448 L----PQEWIQLHYQSAPSQSDVALIRFTNPDTRGVLFPAKLHRSGYI-----TV 494
DB 2200 TYPQPDRIV-----TTPGTAAAPVGNKTP-----IMFASVVRRTGDVNTAGSANGTQ 2249

```

```

QY 495 ANTGRPIVW-----PANGYF-----REDTWVQFYS 521
DB 2250 YGTGSQPLFVTIGLSLNNYSALMPGQPFVWQLTFAS 2286

RESULT 4
COAT_SMSV1
ID COAT_SMSV1 STANDARD; PRT; 702 AA.
AC P36284;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Coat protein (Capsid protein).
OS San Miguel sea lion virus (serotype 1) (SMSV 1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=36406;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=92410750; PubMed=1529644;
RA Neill J.D.;
RT "Nucleotide sequence of the capsid protein gene of two serotypes of
RT San Miguel sea lion virus: identification of conserved and non-
RT conserved amino acid sequences among calicivirus capsid proteins.";
RL Virus Res. 24:211-222(1992).
CC -|- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; M87481; AAA16217.1; -.
CC FIR; A48562; A48562.
CC InterPro; IPR004005; Calici_coat.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DEE3D7 CRC64;

Query Match 10.2%; Score 289; DB 1; Length 702;
Best Local Similarity 25.4%; Pred. No. 5.5e-14;
Matches 128; Conservative 62; Mismatches 190; Indels 124; Gaps 23;

QY 8 AAPNDGAAGLVPEANNETMALEPVAGAS-----IAAPLTGQNNIIDPWIRLNFVQAPNGE 63
DB 151 AESDQFGGADIVTEQGTVQQQVPQAQSALTTLAASTGKTVCDEWTTFFSYHTAVNWS 210

QY 64 FTVSPRNPGEVLLNLELGPENLYLAHLRMYNGVAGGVEQVLLAGNAFTAGKLVFAA 123
DB 211 TT-----EAQKILFGRALSPENLYLAHLRMYNGVAGGVEQVLLAGNAFTAGKLVFAA 266

QY 124 VPPHPLENISPGQITMPPHVIDVRLPVLPLPDVRNPFHYNQNEPRMLVAMLY 183
DB 267 VPP--GIEPVESPTMLQYPHVLFDAQTEPVIPTIDIRKTLVH-SMDDTDTTLLVIMVY 323

QY 184 TPL-RSNGSGDDVFTVSCRVLTRPSPDPFNLYPPTLESKTKPTLPLTIGELTNSRF 242
DB 324 NELINPYQSEPKSCSITVETRPSSDFTSLKLP--GSLKKGSIPLDLP--NSRH 379

QY 243 PVPIDELTYSNESLVQFPQ-----NGRCALDGELOG----- 274
DB 380 --WMGNRWSTIDGFWVQPRVQSNRHFDFDFTTTGWSTFPYVPIEVLTLEKLDRCGQYFK 437

QY 275 -----TTQLLPTAICSFGRINQKVSNGENHVMNQVTNINGTFFP---- 313

```

Db 438 VTDTEKSLVPLGCDGWPDTT--IPTAMTASNGNYDYIVA-----EXRITN-NGTHPKGF 488  
 Qy 314 -----DPTGDPAPLPTGDFSGKLGVLQRDHDNACR-----SHDAVI-- 352  
 Db 489 YIMGNLTTRKVGSDNLGE-----TQOSTRTLFASGVNGYKQONTINPHTKITSLSLVYD 542  
 Qy 353 ATNSAKFTPKLGAIGQITWEEDVHINQP-TKFTFVGLFENEGFNQWTLNPSYGAALTINM 411  
 Db 543 ANNVSAATAK-----TTTWHSTMSHGLVYLVDSPVGSSTKVRVIALPE---AFT-NG 593  
 Qy 412 GLAPPVAPTFPGEQILFFRSHIPL 435  
 Db 594 G-----NFP-----VFFTNKIQI 606

RESULT 5  
 COAT\_FCVF4 STANDARD; PRT; 668 AA.  
 ID COAT\_FCVF4 AC P27405;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS Feline calicivirus (strain Japanese F4) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OX NCBI\_TaxID=11980;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91306470; PubMed=1853578;  
 RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,  
 Miyamura K., Yamazaki S., Mikami T.;  
 RT "Sequence analysis of the 3'-end of feline calicivirus genome.";  
 RL Virology 183:810-814(1991).  
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; D90357; BRA14371.1; -  
 DR PIR; B40481; VCMWFC.  
 DR InterPro; IPR004005; Calici coat.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 668 AA; 73589 MW; 85BBDCEB85804E503 CRC64;  
 Query Match 10.0%; Score 284; DB 1; Length 668;  
 Best Local Similarity 24.8%; Pred. No. 1.2e-13;  
 Matches 122; Conservative 66; Mismatches 194; Indels 110; Gaps 21;  
 Qy 11 SNDGAGLVPEANET--MALEPVPAGASIAAPLTGONNIIDPW-IRLNFVQAPNGEFTVS 67  
 Db 125 ADDGST-TTPEQGTAVGVIAEPASQMSQADMASGKSVDSWEAFPSHISVNW---S 179  
 Qy 68 PRNSPGEVLNLELGPALNLYLAHLSRMVNGYAGGVVQVLLAGNAFTAGKLVFAVDPH 127  
 Db 180 TSETQKLFKQSLGFLNLYLHLSKLVVANSVGSIEVAFSISGSGVFGKLAALVWPP- 238  
 Qy 128 FPLENTSPGQIT---MFPVHVIDVRLVPLPLDVRNFFHYNQNEPRMRLVAMLYT 184  
 Db 239 ----GVDPVQSTSMQYPHVLFDPARQVEPFIPTDLRLSTLHVMSDID-TTSLVIMVYN 293

Qy 185 PL-----RNSGSGDDVFTVSCRVLTREPSPDFDFNYLVPPTLESKTKPFTLPIITIGBELT 238  
 Db 294 DLINPYANDNSSGCIV-----TVETKGPDPFKPHLLKPPG-----SVLTHGSIIP 338  
 Qy 239 NSGRFP---VPIDELYTSPNESLVVQP-----ONGRCALDGELOG--TTQLLPALCAISFRG 288  
 Db 339 SDLIPKSSSLWIGNRYWTDITDFVIRPFVFOANRHFDFENQETAGWSTPRFRPITI----- 393  
 Qy 289 RINQKVGSGENHWNQVYTNINGTP-----FDPTGDPVAPLG----- 324  
 Db 394 TISEKNGSKLGIQVATDYIIPGDPGWPTTTADKLIIPAGDYSITTGENDIKTAQAYDT 453  
 Qy 325 -----TPDFSGKLGVLQRDHDNACRSHDAVIAT-----NSAKFTPKLGAIGQITWEE 373  
 Db 454 AAVVKNYTNFRGMVYICSLQRAWGDKKISNTAFITTAIRDGNEIKPSNTIDMTKLAVYO- 512  
 Qy 374 DVHINOPTK-----FTFVGLFENEGFNQWTLNPSYGAALTINMGLAPPVAPTFPG 423  
 Db 513 -DTHVEQEVQTSDDTLALLGYTGIG-EEAIGSNRDRVRIS-----VLPAGARGG 561  
 Qy 424 EQILFFRSHIPL 435  
 Db 562 NHPIFYKNISKL 573

RESULT 6  
 COAT\_FCVF9 STANDARD; PRT; 671 AA.  
 ID COAT\_FCVF9 AC P27406;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Coat protein (Capsid protein).  
 GN CP276.  
 OS Feline calicivirus (strain F9) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OX NCBI\_TaxID=11981;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92410623; PubMed=1529544;  
 RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,  
 Turner P.C.;  
 RA "The complete nucleotide sequence of a feline calicivirus";  
 RT Virology 190:443-448(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92339673; PubMed=1633955;  
 RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;  
 RA "Cloning and sequence determination of the feline calicivirus strain  
 F9.";  
 RT Biochem. Soc. Trans. 20:268-268(1992).  
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; M86379; AAA79327.1; -  
 DR EMBL; Z11536; CAA77636.1; -  
 DR PIR; B43382; VCMWF9.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 671 AA; 73441 MW; 33BEE86D8370D5E5 CRC64;

Query Match  
 Best Local Similarity 9.8%; Score 277.5; DB 1; Length 671;  
 Matches 76; Conservative 38; Mismatches 97; Indels 31; Gaps 9;

QY 30 EPVAGASIAAPITGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPGEVLLNLGLPELNPY 88  
 DB 145 EPSAQWSTAADWATGKSDVSEWEPFSTSVNW-----STSETQKILFKQSLGPLNPY 200  
 QY 89 LAHLGRMYNGYAGGVEVQVLLAGNAFTAGKLVFAAPVPHFPLENISPGQIT---MPHVI 145  
 DB 201 LEHLAKLYVWAGSIEVRSISGSGVFGKLAIVVPP-----GVDPPQSTSMQLQPHVL 255  
 QY 146 IDVRLTEPVLPLPVRNNFFHYNQNEPRMELVAMLYTPLRSNGSDVFTVSC--RVL 203  
 DB 256 FPARQVEPVIFCLPDLRSTLYHLMSTD-TTSLVIMVNDL-INPYANDANGSGGIIVTE 313  
 QY 204 TRPSPDFDNVLPPTLESKTPFTLPILTIGELTNSRPP-----VPIDELYTSPNESLVV 259  
 DB 314 TRPGPDFKFLKPPG-----SMLTHGSIPLDLIPKTSLSLWIGNRYWSDITDFVI 363  
 QY 260 QP 261  
 DB 364 RP 365

RESULT 7  
 COAT\_SMSV4  
 ID COAT SMSV4 STANDARD; PRT; 703 AA.  
 AC P36285;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS San Miguel sea lion virus (serotype 4) (SMSV 4).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OX NCBI\_TaxID=36407;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92410750; PubMed=1529644;  
 RA Neill J.D.;  
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of  
 RT San Miguel sea lion virus: identification of conserved and non-  
 RT conserved amino acid sequences among calicivirus capsid proteins.";  
 RL Virus Res. 24:211-222(1992).  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M87482; AAA16220.1; -.  
 DR PIR; C48562; C48562.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 482 482 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 703 AA; 77721 MW; C5DAD8223B261073 CRC64;  
 Query Match  
 9.6%; Score 273.5; DB 1; Length 703;

Best Local Similarity 28.8%; Pred. No. 8e-13;  
 Matches 97; Conservative 51; Mismatches 150; Indels 39; Gaps 15;

QY 8 AAPSDNGAAGLVPEANNETM-----ALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAPNGE 63  
 DB 151 AESDQFGSAEIVTEGQITVQOOPAPATLATATATGK-SVEQEWTFPSYHTSINW 209  
 QY 64 FTVSPRNSPGEVLLNLGLPELNPYLAHLGRMYNGYAGGVEVQVLLAGNAFTAGKLVFAA 123  
 DB 210 STV-----ESQKILYQALNPSINPYLDHIAKLYSTWSGIDVRFTVSGSGVFGKLAALL 266  
 QY 124 VPPHF-PLENISPGQITMPHVIIDVRLTEPVLPLPVRNNFFHYNQNEPRMELVAML 182  
 DB 267 VPPGVEPIESVS---MLQYPHVLFARQTEPVIFTPIDIRKTLFHSMDTD-TTKLVINP 322  
 QY 183 YTPLRSNGSDVFTVSCRVLTRPSPDFDNVLPPTLESKTPFTLPILTIGELTNSRF 242  
 DB 323 Y-----ENGV-ENKTCSTIVETRPADTFALLKPP--GSLIKHSIP-----SDILPRNS 371  
 QY 243 VPVIDELYTSPNESLVVQFQ-----NGRCALDGELOG--TTQLLPTAICSGRGRINQKVG 296  
 DB 372 AHMGNRMWSTISGFSVQPRVFQSNRHFDFDSTTTGWTSPYVPIEI-----KIQGKVG 426  
 QY 297 EN---HVNMQVTNINGTTPFDPDGDVPAPLGTPDFSG 330  
 DB 427 NNMWFHVIDTDKALVPGIP-DGWPDTTIPDETATNG 462

RESULT 8  
 COAT\_FVCV6  
 ID COAT FVCV6 STANDARD; PRT; 668 AA.  
 AC P27404;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OX NCBI\_TaxID=11979;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=91374597; PubMed=1716692;  
 RA Neill J.D.; Reardon I.M.; Heinrichson R.L.;  
 RT "Nucleotide sequence and expression of the capsid protein gene of  
 RT feline calicivirus".  
 RL J. Virol. 65:5440-5447(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Neill J.D.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 CC -!- PTM: The N-terminus is blocked.  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; M32819; AAA42925.1; -.  
 DR EMBL; U13992; AAC13993.1; -.  
 DR PIR; A40507; VCMWFF.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).

FT CARBOHYD 459 459 N-LINKED (GLNAC. . .) (POTENTIAL).  
FT CARBOHYD 615 615 N-LINKED (GLNAC. . .) (POTENTIAL).  
SQ SEQUENCE 668 AA; 73550 MW; 9E52312108D441 CRC64;  
  
Query Match 9.6%; Score 272.5; DB 1; Length 668;  
Best Local Similarity 31.3%; Pred. No. 8.9e-13;  
Matches 77; Conservative 37; Mismatches 93; Indels 39; Gaps 9;  
  
QY 30 EPVAGASIAAPLTGQNNIIDPW-IRLNFVQAPNGEFTVSPRNSPGEVLLNLELGPENPY 88  
DB 145 EFNAQWSTADWATKGVSDSEWAFSHTSVNW-----STSETQKILFKOSLGLPLNPY 200  
QY 89 LAHLSRMYNGYAGVEVQVLLAGNAFTAGKLVFAAVPHFPLENISPGQIT---MPPHVI 145  
DB 201 LTHLAKLYAVSGVSDVRSISGSGVFGKLAIVVPP-----GIDPVQSTMLQYPHVL 255  
QY 146 IDVRLTEPVLPLPDPVRNFFHYNQNEPRMLVAMLYTP-----RSNSGDDVFTVS 199  
DB 256 FDAQVEPVIFSPDLRSTLYHLSMDTD-TTSLVIMVYNDLINPYANDNSGCVI----- 310  
QY 200 CRVLTSPDPDFNLYVPPTLESKTPFTLPILTIGELTNSRPP-----VPIDELYTSPNE 255  
DB 311 -TVETKPGDPKFKHLKPPG-----SMLTHGSPFSDLPKSSSLWIGNRFSWSDIT 359  
QY 256 SLVVQP 261  
DB 360 DFVIRP 365  
  
RESULT 9  
POLN MANCV STANDARD; PRT; 2208 AA.  
ID -POLN MANCV  
AC Q69014;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polyprotein (Contains RNA-dinect RNA polymerase  
DE (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like  
DE protein); Coat protein).  
OS Manchester virus (Human enteric calicivirus).  
OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Sapovirus.  
OX NCBI\_TaxID=82659;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95390791; PubMed=7661689;  
RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;  
RT "Human enteric caliciviruses have a unique genome structure and are  
RT distinct from the Norwalk-like viruses."  
RL Arch. Virol. 140:1345-1356(1995).  
CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).  
CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR  
CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY  
CC SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate + N diphosphate +  
CC {RNA} (N).  
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; X86560; CAA60262.1; --  
DR MEROPS; C24.0PW; --  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR004004; Calici\_pol\_hel.  
DR InterPro; IPR000317; Peptidase\_C24.

DR InterPro; IPR000605; RNA\_helicase.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
DR Pfam; PF03510; Peptidase\_C24; 1.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
DR PRINTS; PR00918; CALICIVIRUSNS.  
DR POLYPROTEIN; Transferase; RNA-directed RNA polymerase; Hydrolase;  
KW Thiol protease; Helicase; ATP-binding; Coat protein.  
FT CHAIN ? ? HELICASE (P2C-LIKE).  
FT CHAIN ? ? PROTEASE (P3C).  
FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.  
FT NP\_BIND 408 415 ATP (POTENTIAL).  
FT ACT\_SITE 1097 1097 PROTEASE (BY SIMILARITY).  
FT ACT\_SITE 1112 1112 PROTEASE (BY SIMILARITY).  
SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;  
  
Query Match 9.2%; Score 262; DB 1; Length 2208;  
Best Local Similarity 21.0%; Pred. No. 2.8e-11;  
Matches 133; Conservative 73; Mismatches 251; Indels 176; Gaps 23;  
  
QY 2 KMASNDAAAPSGNDGAAGLVPEANNETWALEPFAVAGASIAAPLTGQNNIIDPWIRLNFVQAPN 61  
DB 1647 EMEGNSNP-----PKQSNPNVVDP---PGTTGPTTSHVVVAVP-----EQPN 1688  
QY 62 G-----EFTVSPRNSPGEVLLNLELGPENPYLAH 91  
DB 1689 GAAQRLLEAVATGAIQSNVPEAIRNCFAVETFAWDRMPTGFLGISLSHPNINPTSH 1748  
QY 92 LSRMYNGYAGVEVQVLLAGNAFTAGKLVFAAVPHFPLENISPGQIT---MPPHVIDV 148  
DB 1749 LSGWAGWGGSGFEVRLSISGSGVFAAGRIASVIPP-----GVDPSSIRDPGVLPFAFVDA 1803  
QY 149 RTLEPVLPLPDPVRNFFHYNQNEPRMLVAMLYTPLSNSGDDVFTVSCRVLRPSP 208  
DB 1804 RITEPVSFMIEDVRAVDYHRMDGAEPTCSLGFVYQPLLPFPSTTAVSTCWVSEYKPGG 1863  
QY 209 DFDENLYVP---TLESKTKPFTLPILTIGELTNSRFPVPIDELYTSPNESLVVQPNQGR 266  
DB 1864 DFDCLLPQPGQOMENGVSPEGILLPRLGLYSRGNRVGLVGVMLVAEHQVNRHFNNS 1923  
QY 267 ALDGELOQTQLLPTAICSPFRINQKVSNGENHVNMQVNTINGTTPDPTGDVPAIATGP 326  
DB 1924 VTFGWSAPVNPMAAEIVT---NOAHSTSRHAW-LSIGAQNKGPLFP-----GIP 1969  
QY 327 DFGSKGLFVLSQRDHDNACRSHDAVIATNSAKFTPKLCAI---QIGTWEEDDVHINQP--- 381  
DB 1970 N-----HFPDSCASTVGMMDTSLGRRPSTGCGPAISFQNNNGDVYENDTSPV 2017  
QY 382 -----TKFTPVGL-----FENEGFN-----QWTLPNYSGLAT 408  
DB 2018 MFATYDPLTSGTGVALTNSINPASLALVRISNNDFTSGFANDKNVVQVMSWMTYGNQ 2077  
QY 409 LNMGLAPPVAPT-----FPGEQILPFRSHIPLKGGVADPVIDCL 447  
DB 2078 IR-GQVTPMSGNTYFTTGTANTLVLMQERMLSYDGHQAILYSSQ----- 2122  
QY 448 LPOEWIQLHYQESAPQSQSDVALIRFNTDPTGRVLFPAKLRHSGYITVANTGSRPIVSPAN 507  
DB 2123 ---ERTAEYFQNDIVNIPENSMAVF-NVETNSASFOIGIRPDGVMYTG--GSIGNVPLE 2176  
QY 508 GYFRFDPTWNQFYSLAPMG-----TGN-GRRR 533  
DB 2177 PETRF-----QYVGILPLSAALSGPSGNMGRAR 2204  
  
RESULT 10  
POLG\_CX16T



ID AC 9QF31; POLG\_CX16T STANDARD; PRT; 2193 AA.

DT 10-OCT-2003 (Rel. 42, Created)

DT 10-OCT-2003 (Rel. 42, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2 (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].

DE Cocksackievirus A16 (strain Tainan/5079/98).

OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.

OC NCBI\_TaxID=231417;

OX [1]

RN SEQUENCE FROM N.A.

RX MEDLINE=21427116; PubMed=11536241;

RA Yan J.-J., Su I.-J., Chen P.-F., Liu C.-C., Yu C.-K., Wang J.-R.;

RT "Complete genome analysis of enterovirus 71 isolated from an outbreak in Taiwan and rapid identification of enterovirus 71 and coxsackievirus A16 by RT-PCR.";

RT J. Med. Virol. 65:331-339 (2001).

RL

CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular membranes and is associated with viral RNA synthesis.

CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.

CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.

CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA) (N).

CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.

CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.

CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL; AF177911; AAD55085.1; ..

DR HSP; P03300; IPOV.

DR InterPro; IPR003593; AAA ATPase.

DR InterPro; IPR004004; Calici\_pol\_hel.

DR InterPro; IPR009003; Cys Ser trypsin.

DR InterPro; IPR000199; Pept\_3C\_picorn.

DR InterPro; IPR000081; Peptidase\_C3.

DR InterPro; IPR003138; Pico\_P1A.

DR InterPro; IPR002527; Pico\_P2B.

DR InterPro; IPR001676; Rhv.

DR InterPro; IPR006005; RNA helicase.

DR InterPro; IPR007095; RNA\_pol\_DS\_PS.

DR InterPro; IPR001205; RNA\_pol\_P3D.

DR InterPro; IPR007094; RNA\_pol\_PSVir.

DR InterPro; IPR008975; Viral\_cap\_coat.

DR Pfam; PF00548; Cys-protease-3C; 1.

DR Pfam; PF02226; Pico\_P1A; 1.

DR Pfam; PF00947; Pico\_P2A; 1.

DR Pfam; PF01552; Pico\_P2B; 1.

DR Pfam; PF00073; rhv\_3.

DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.

DR Pfam; PF00910; RNA\_helicase; 1.

DR PRINTS; PR00918; CALICIVIRUSNS.

DR ProDom; PD001125; Cys protease\_3C; 1.

DR ProDom; PD001306; Pico\_P2A; 1.

DR ProDom; PD001274; Pico\_P2B; 1.

DR SMART; SM00382; AAA; 1.

DR Polyprotein; Coat protein; Core protein; Transferase;

KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;

KW Lipoprotein. 2 69 COAT PROTEIN VP4.

FT CHAIN 70 323 COAT PROTEIN VP2.

FT CHAIN 324 565 COAT PROTEIN VP3.

FT CHAIN 566 862 COAT PROTEIN VP1.

FT CHAIN 863 1012 CORE PROTEIN P2A.

FT CHAIN 1013 1111 CORE PROTEIN P2B.

FT CHAIN 1112 1440 CORE PROTEIN P2C.

FT CHAIN 1441 1526 CORE PROTEIN P3A.

FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VPG.

FT CHAIN 1549 1731 PICORNAIN 3C.

FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.

FT LIPID 2 N-myristoyl glycine (in host) (By similarity).

FT ACT SITE 1695 PROTEASE (POTENTIAL).

FT ACT SITE 1709 PROTEASE (POTENTIAL).

SQ SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;

Query Match 5.0%; Score 142; DB 1; Length 2193;

Best Local Similarity 22.7%; Pred. No. 0.028;

Matches 127; Conservative 60; Mismatches 223; Indels 150; Gaps 31;

QY 1 MKMASNDAAPSNDG-AAGLVP-----EANNETMALEPVAGASIAAPLTGQNNI 47

DB 329 LKPGTNQFLTTDDGVSAFILPGFHPTPIHIPGEVHN---LLEICRVETILEVNNLKTN 385

QY 48 IDPMRLNLF---VOAPNGSEFTVSPRNSPGEVLLNLDELGPENLPYLAHLSRMVNGVAGVE 104

DB 386 TTFMQRLCFPVSVQSKTGBELCAAFADPFORD-----GFWQSTILGQLCRYTQSGSLE 439

QY 105 VOVLAGNAFTAGKLVAFAVPPHPPLENISPGQIT--MPHHVIDVRLTEPLVLLPDVR 162

DB 440 VTFMAGSFMAITGKMLIAYTPPG---GNVPADRIATMLGTHVIMDFGLQSSVTLVVPWIS 496

QY 163 NNFHYNQNEPRMLVAMLYTPLRSNGSGDDVFTVSCRVLTRPSDFDPFNVLPPTLES 222

DB 497 NT--HY-----RAHAR-AGYFDY-----511

QY 223 KTKPPTLITIGELTNSRFPVPIDELYTSNPSLVVQPN---GRCALDGLSGTQQL 279

DB 512 ----YTTGLIIWYQTN--YVVPICAPTAYIVALAAQDNFTMKLCKDTEIEQTANIQ 565

QY 280 PTAICSFGRINQKVSNGENHVMQVNTINGTPDPDGDVPAP---LGT---PDFSGKLF 333

DB 566 GDPIADM---IDQTV---NNQVNRSLTALQVLP--TAADTEASSHRLGTGVVPALQAABT 617

QY 334 GVLSDRDHNDACR-----SHDAVIATNSAKFTPKLGAIQIGTWEEDDVHINOPTKTPVG 388

DB 618 GASSNASDKNLTETRCVLNHHSTQETAIGNFPFSRAGLSI-----ITWPT---TG 664

QY 389 LFENEGFNQWTLPNYSGALTINMGLAPPVAPTFPGEQILFPRSHIPLKGGVADPVIDCLL 448

DB 665 QNTDGYNNWDI-DLMGVAQLRRKC-----ELFTYKRFDAEFVFAKPNGE-LV 712

QY 449 PQEWIQLHY-QESAP---SQSDVALIRFTNPDTRGVLFPEAKLHRSGYITVANTGSRPIVV 504

DB 713 PQ-LIQYVYVPPGAPKPTSRDSFAQTATNPS-----VFVKMTDPPPAQVSV-----PFMS 761

QY 505 PANGYFRFDTWVNPQYSLAP 524

DB 762 PASAY-----QW---FYDGY 774

## RESULT 11

POLG\_CX16G

ID POLG\_CX16G STANDARD; PRT; 2193 AA.

AC Q65900;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2 (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Core protein



RN SEQUENCE FROM N.A.  
 RX MEDLINE=87010550; PubMed=3020156;  
 RA Hughes P.J., Evans D.M.A., Minor P.D., Schild G.C., Almond J.W.,  
 RA Stanway G.;  
 RT "The nucleotide sequence of a type 3 poliovirus isolated during a  
 RT recent outbreak of poliomyelitis in Finland.";  
 RL J. Gen. Virol. 67:2093-2102(1986).  
 CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
 CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-I-Gly bond in the  
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
 CC substituted for Gln, and Ser or Thr for Gly.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC [RNA] (N).  
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X04468; CAA28155.1; -;  
 DR PIR; A27245; GNNY27.  
 DR HSSP; Q84790; LPVC.  
 DR MEROPS; C03.001; -;  
 DR MEROPS; C03.020; -;  
 DR InterPro; IPR003593; AAA\_ATPase.  
 DR InterPro; IPR004004; Calci\_pol\_hel.  
 DR InterPro; IPR009003; Cys\_ser\_trypsin.  
 DR InterPro; IPR000199; Pept\_3C\_picorn.  
 DR InterPro; IPR000081; Peptidase\_C3.  
 DR InterPro; IPR003138; Pico\_P1A.  
 DR InterPro; IPR002527; Pico\_P2B.  
 DR InterPro; IPR001676; Rhv.  
 DR InterPro; IPR006005; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_P5vir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00548; Cys-protease-3C; 1.  
 DR Pfam; PF02226; Pico\_P1A; 1.  
 DR Pfam; PF00947; Pico\_P2A; 1.  
 DR Pfam; PF01552; Pico\_P2B; 1.  
 DR Pfam; PF00073; rhv\_3;  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PR00918; CALICVIRUSNS.  
 DR ProDom; PD001125; Cys\_protease\_3C; 1.  
 DR ProDom; PD001306; Pico\_P2A; 1.  
 DR ProDom; PD001274; Pico\_P2B; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR Polyprotein; Coat protein; Core protein; Transferrase;  
 KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 KW Lipoprotein.  
 FT CHAIN 2 69 COAT PROTEIN VP4.  
 FT CHAIN 70 340 COAT PROTEIN VP2.  
 FT CHAIN 341 578 COAT PROTEIN VP1.  
 FT CHAIN 579 878 COAT PROTEIN VP3.  
 FT CHAIN 879 1027 CORE PROTEIN P2A.  
 FT CHAIN 1028 1124 CORE PROTEIN P2B.  
 FT CHAIN 1125 1453 CORE PROTEIN P2C.  
 FT CHAIN 1454 1540 CORE PROTEIN P3A.  
 FT CHAIN 1541 1562 GENOME-LINKED PROTEIN VPg.  
 FT CHAIN 1563 1745 PICORNAIN 3C.  
 FT CHAIN 1746 2206 RNA-DIRECTED RNA POLYMERASE P3D.

FT LIPID 2 2 N-myristoyl glycine (in host) (By  
 FT similarity).  
 FT ACT SITE 1709 1709 PROTEASE (POTENTIAL).  
 FT ACT SITE 1723 1723 PROTEASE (POTENTIAL).  
 SQ SEQUENCE 2206 AA; 245731 MW; F226AD85403C37BA CRC64;  
 Query Match 4.6%; Score 130; DB 1; Length 2206;  
 Best Local Similarity 18.7%; Pred. No. 0.23;  
 Matches 114; Conservative 71; Mismatches 178; Indels 248; Gaps 29;  
 QY 11 SNDGAAGLVLP-----EANNETMALEBVA-----GASIAAPLT----- 42  
 DB 270 TNSATLVLPYNALSIDSMWKNWNGIALPLSLDFDAQSSVIRPITVTIAPMCSEPN 329  
 QY 43 GQNNIIDPWIR-LNFVQAPNG-----EFTVSPR-NSPGEV----- 75  
 DB 330 GLRNVTA PKLGLPVLNTPGSGNOYLTSNHSQSPCAIPFEDVTPPIDIPGEVKNVMELAEI 389  
 QY 76 -----LNLELGPENLPYLH-----LSRMVNG 98  
 DB 390 DTMIPLNLENTKNTMDMYRVLSDSANLSGPICLSLSPAADPRLSHTMLGEVLNYTH 449  
 QY 99 VAGGVEVQVLLAGNAFTAGKLVFAAVPPHFPLENISPGQITWPHVIDVRLTLEPVLPL 158  
 DB 450 WAGSLKFTFLFCGSMWATGKLLVAYAPG-AQPPTSRSKEMLGTHVINDLGLQSSCTMVV 508  
 QY 159 PDVRNFFHYNQNE-----PRMLVAMLYTLRSNGSGDDVFTVSCRVLTRPSP 208  
 DB 509 PWISNVTYRQTQDSFTGGYISMFYQTRIVVPLSTPKAMDMLG---FVSACN----- 558  
 QY 209 DFDFNYLVPPTLESKTFTLPILTIGBLTNSRFPVPIDELYTSNESLVV-----OPQNG 264  
 DB 559 DFSVLLRDIT-----HISQAAMPQGVDDLITEVAQNALALSPLKPSQN 602  
 QY 265 RCALDGBELQGTQLLPTALCSFRGRINOKVSGENHVMQVNTINGTPDPTGDVPAPLG 324  
 DB 603 -----DTRKSGPAH--SKREVTLTAVETGATN-----PL- 631  
 QY 325 TPDFSGKLGVLQRDHD-----NACRSHDAVI-----ATNSAKFTPKLGAIQ 367  
 DB 632 VPSDVTQTRHVIQQRSRSESTIESPFARGAC---VALIEVDNEQPATNVQKL----- 680  
 QY 368 IGTWE---EDDVHINQPTKTPVGLFENEGNQWTLPNVSGALTLMGLA----- 414  
 DB 681 FATWRTYTKDTVQLRRKLEFFTYSRFDME-FTFVVVTANFTNS---NNGHALNQVYQIMVI 736  
 QY 415 PVVAPTFPGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHYQES-----APS 463  
 DB 737 PPGAPT-----PRSWDDYTWQTSNPSPIFYTYGAAPA 768  
 QY 464 QSDVALIRFTN 474  
 DB 769 RISVFPVGLAN 779  
 RESULT 13  
 POLG HE701  
 ID POLG HE701 STANDARD; PRT; 2194 AA.  
 AC P32537;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2  
 DE (P1B); Coat protein VP3 (P1C); Coat protein VP4 (P1A); Coat protein VP2  
 DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
 DE linked protein VPg (P3B); Picornain 3C (SC 3.4.22.28) (Protease 3C)  
 DE (P3C); RNA-directed RNA polymerase P3D (SC 3.4.22.28) (Protease 3C)].  
 OS Human enterovirus 70 (strain J670/71) (EV 70).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 OX NCBI\_TaxID=31915;  
 RN [1]  
 RP SEQUENCE FROM N.A.

FT ACT\_SITE 1701 1701 PROTEASE (POTENTIAL).  
 FT ACT\_SITE 1715 1715 PROTEASE (POTENTIAL).  
 SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;  
 Query Match 4.5%; Score 128; DB 1; Length 2194;  
 Best Local Similarity 21.0%; Pred. No. 0.32;  
 Matches 61; Conservative 51; Mismatches 104; Indels 74; Gaps 14;  
 QY 48 IDPWIRLNFQAPNG-----EFTVSPRNSPGEVLLNLEL-----GPELNPYLALHLSRMVNG 98  
 DB 368 IESMEINNVNDASGVERLRVQISQSDMDQLFNIPLDIQLEGLPLRLTLGNISRYTH 427  
 QY 99 YAGGEVQVILLAGNAFTAGKLVFAAVPP--HFFLENISPGQITMFFHVIDVRLPEVLL 156  
 DB 428 WSGSLEWTFMFCGSMFTTGLIICVTPGSGSPTDRM---QAMLAHVVMDFGLOSIIT 484  
 QY 157 PLPDVRNPNFHYNQNEPRML---VAMLYTPLRN-----GSGDDVFVTSKRV----- 202  
 DB 485 IIPWISGS--HYRMFNTDAKAINANVGYVTCFMQTNLVPVGAADOCYIVGMVAACKDFN 542  
 QY 203 --LTRPSPDFDNVLPVPP-----TLESKTKPF-----TLPIITIGELTN 239  
 DB 543 LRLMRSDPDIGQSAIIPQEAATQIGEVKTVANTVESEIKAEGLVIPSINAVETGATSN 602  
 QY 240 SRFPVPIDELYTSPNSLVVQPNQRCALDGLQCTQLLPTAICSFRGR 289  
 DB 603 -----TPPEAI-----QTRTVIN--MHGTAECLE---VENFLGR 631  
 RESULT 14  
 ID POLG\_FMDV1 STANDARD; PRT; 2333 AA.  
 AC P03306; Q64768; Q84750; Q84751; Q84752; Q84753; Q84754; Q84760;  
 AC Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;  
 AC Q84769; Q89824;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DE 10-OCT-2003 (Rel. 01, Last sequence update)  
 DE Genome polypeptide [Contains: Nonstructural protein P20A; Coat  
 proteins VP1 TO VP4; Core protein P52; Genome-linked proteins VPg1 TO  
 VPg3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed  
 RNA polymerase P56A (EC 2.7.7.48)].  
 OS Foot-and-mouth disease virus (strain A10-61) (Aphthovirus A) (FMDV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Aphthovirus.  
 OX NCBI\_TaxID=12112;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84169547; PubMed=6324120;  
 RA Carroll A.R., Rowlands D.J., Clarke B.E.;  
 RT "The complete nucleotide sequence of the RNA coding for the primary  
 translation product of foot and mouth disease virus.";  
 RL Nucleic Acids Res. 12:2461-2472(1984).  
 RN [2]  
 RP SEQUENCE OF 115-1048 FROM N.A.  
 RX MEDLINE=82211814; PubMed=6282711;  
 RA Boothroyd J.C., Harris T.J.R., Rowlands D.J., Lowe P.A.;  
 RT "The nucleotide sequence of cDNA coding for the structural proteins  
 of foot-and-mouth disease virus.";  
 RL Gene 17:153-161(1982).  
 CC -I- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the  
 poliovirus polypeptide. In other picornavirus reactions Glu may be  
 substituted for Gln, and Ser or Thr for Gly.  
 CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 {RNA} (N).  
 CC -I- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 each of which is composed of one copy each of proteins VP1, VP2,  
 VP3, and VP4.  
 CC -I- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: D00820; BAA18891.1; -;  
 CC PIR: A36253; GNNYE7.  
 CC HSP: P03300; IPOV.  
 CC MEROPS: C03.00A; -;  
 CC MEROPS: C03.00B; -;  
 CC InterPro: IPR003593; AAA\_ATPase.  
 CC InterPro: IPR004004; Calici\_pol\_hel.  
 CC InterPro: IPR005003; Cys\_Ser\_trypsin.  
 CC InterPro: IPR000199; Pept\_3C\_picorn.  
 CC InterPro: IPR000081; Peptidase\_C3.  
 CC InterPro: IPR003138; Pico\_P1A.  
 CC InterPro: IPR002527; Pico\_P2B.  
 CC InterPro: IPR001676; Rhv.  
 CC InterPro: IPR000605; RNA\_helicase.  
 CC InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 CC InterPro: IPR001205; RNA\_pol\_P3D.  
 CC InterPro: IPR007094; RNA\_pol\_PSVir.  
 CC InterPro: IPR008975; Viral\_cap\_coat.  
 CC Pfam: PF00548; Cys-protease-3C; 1.  
 CC Pfam: PF02226; Pico\_P1A; 1.  
 CC Pfam: PF00947; Pico\_P2A; 1.  
 CC Pfam: PF01552; Pico\_P2B; 1.  
 CC Pfam: PF00073; rhv; 3.  
 CC Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
 CC Pfam: PF00910; RNA\_helicase; 1.  
 CC PRINTS: PR00918; CALICIVIRUSNS.  
 CC ProDom: PD001125; Cys\_protease\_3C; 1.  
 CC ProDom: PD001306; Pico\_P2A; 1.  
 CC ProDom: PD001274; Pico\_P2B; 1.  
 CC SMART; SM00382; AAA; 1.  
 CC RNA-directed RNA polymerase; Core protein; Transferase;  
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 CC Lipoprotein.  
 CC CHAIN 2 69 COAT PROTEIN VP4.  
 CC CHAIN 70 319 COAT PROTEIN VP2.  
 CC CHAIN 320 561 COAT PROTEIN VP3.  
 CC CHAIN 562 871 COAT PROTEIN VP1.  
 CC CHAIN 872 1014 CORE PROTEIN P2A.  
 CC CHAIN 1015 1113 CORE PROTEIN P2B.  
 CC CHAIN 1114 1443 CORE PROTEIN P2C.  
 CC CHAIN 1444 1532 CORE PROTEIN P3A.  
 CC CHAIN 1533 1554 GENOME-LINKED PROTEIN VPg.  
 CC CHAIN 1555 1737 PICORNAIN 3C.  
 CC CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.  
 CC CHAIN 2 N-myristoyl glycine (in host) (By  
 CC LIPID 2 similarity).

RA MEDLINE=91037960; PubMed=2172447;  
 RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D.,  
 RA Minor P.D., Almond J.W.;  
 RT "The complete nucleotide sequence of enterovirus type 70:  
 relationships with other members of the picornaviridae.";  
 RL J. Gen. Virol. 71:2291-2299(1990).  
 CC -I- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
 O/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.  
 CC -I- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the  
 poliovirus polypeptide. In other picornavirus reactions Glu may be  
 substituted for Gln, and Ser or Thr for Gly.  
 CC -I- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 {RNA} (N).  
 CC -I- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 each of which is composed of one copy each of proteins VP1, VP2,  
 VP3, and VP4.  
 CC -I- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -I- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: D00820; BAA18891.1; -;  
 CC PIR: A36253; GNNYE7.  
 CC HSP: P03300; IPOV.  
 CC MEROPS: C03.00A; -;  
 CC MEROPS: C03.00B; -;  
 CC InterPro: IPR003593; AAA\_ATPase.  
 CC InterPro: IPR004004; Calici\_pol\_hel.  
 CC InterPro: IPR005003; Cys\_Ser\_trypsin.  
 CC InterPro: IPR000199; Pept\_3C\_picorn.  
 CC InterPro: IPR000081; Peptidase\_C3.  
 CC InterPro: IPR003138; Pico\_P1A.  
 CC InterPro: IPR002527; Pico\_P2B.  
 CC InterPro: IPR001676; Rhv.  
 CC InterPro: IPR000605; RNA\_helicase.  
 CC InterPro: IPR007095; RNA\_pol\_DS\_PS.  
 CC InterPro: IPR001205; RNA\_pol\_P3D.  
 CC InterPro: IPR007094; RNA\_pol\_PSVir.  
 CC InterPro: IPR008975; Viral\_cap\_coat.  
 CC Pfam: PF00548; Cys-protease-3C; 1.  
 CC Pfam: PF02226; Pico\_P1A; 1.  
 CC Pfam: PF00947; Pico\_P2A; 1.  
 CC Pfam: PF01552; Pico\_P2B; 1.  
 CC Pfam: PF00073; rhv; 3.  
 CC Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
 CC Pfam: PF00910; RNA\_helicase; 1.  
 CC PRINTS: PR00918; CALICIVIRUSNS.  
 CC ProDom: PD001125; Cys\_protease\_3C; 1.  
 CC ProDom: PD001306; Pico\_P2A; 1.  
 CC ProDom: PD001274; Pico\_P2B; 1.  
 CC SMART; SM00382; AAA; 1.  
 CC RNA-directed RNA polymerase; Core protein; Transferase;  
 CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
 CC Lipoprotein.  
 CC CHAIN 2 69 COAT PROTEIN VP4.  
 CC CHAIN 70 319 COAT PROTEIN VP2.  
 CC CHAIN 320 561 COAT PROTEIN VP3.  
 CC CHAIN 562 871 COAT PROTEIN VP1.  
 CC CHAIN 872 1014 CORE PROTEIN P2A.  
 CC CHAIN 1015 1113 CORE PROTEIN P2B.  
 CC CHAIN 1114 1443 CORE PROTEIN P2C.  
 CC CHAIN 1444 1532 CORE PROTEIN P3A.  
 CC CHAIN 1533 1554 GENOME-LINKED PROTEIN VPg.  
 CC CHAIN 1555 1737 PICORNAIN 3C.  
 CC CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.  
 CC CHAIN 2 N-myristoyl glycine (in host) (By  
 CC LIPID 2 similarity).

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC	EMBL; V011330; CAA24361.1; -	
DR	EMBL; X00429; CAA25127.1; -	
DR	MEROPS; C03.008; -	
DR	MEROPS; C28.001; -	
DR	InterPro; IPR004004; Calici_pol_hel.	
DR	InterPro; IPR009003; Cys_Ser_ttrypsin.	
DR	InterPro; IPR004080; FMDVPicoat.	
DR	InterPro; IPR008739; Peptidase_C28.	
DR	InterPro; IPR001676; Rhv.	
DR	InterPro; IPR000605; RNA_helicase.	
DR	InterPro; IPR007095; RNA_pol_DS_Ps.	
DR	InterPro; IPR001205; RNA_pol_P3D.	
DR	InterPro; IPR007094; RNA_pol_PSVir.	
DR	InterPro; IPR008975; Viral_cap_coat.	
DR	Pfam; PF05408; Peptidase_C28; 1.	
DR	Pfam; PF00073; rhv; 3.	
DR	Pfam; PF00680; RNA_dep_RNA_pol; 1.	
DR	Pfam; PF00910; RNA_helicase; 1.	
DR	PRINTS; PR00918; CALICIVIRUSNS.	
DR	PRINTS; PR01542; FMDVPICOAT.	
KW	Polyprotein; Coat protein; RNA-directed RNA polymerase;	
KW	Transferase; Hydrolase; Thiol protease; Nonstructural protein;	
KW	Myristate; Lipoprotein.	
FT	CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.	
FT	CHAIN 202 286 COAT PROTEIN VP4.	
FT	CHAIN 287 504 COAT PROTEIN VP2.	
FT	CHAIN 505 725 COAT PROTEIN VP3.	
FT	CHAIN 726 937 COAT PROTEIN VP1.	
FT	CHAIN 938 1578 CORE PROTEIN P52.	
FT	CHAIN 1579 1601 GENOME-LINKED PROTEIN VPGL.	
FT	CHAIN 1602 1625 GENOME-LINKED PROTEIN VPGL.	
FT	CHAIN 1626 1649 GENOME-LINKED PROTEIN VPGL.	
FT	CHAIN 1650 1863 PROTEASE P20B.	
FT	CHAIN 1864 2333 RNA-DIRECTED RNA POLYMERASE P56A.	
FT	LIPID 202 202 N-myristoyl glycine (in host).	
FT	CONFLICT 396 396 S -> C (IN REF. 2).	
FT	CONFLICT 632 632 P -> L (IN REF. 2).	
SQ	SEQUENCE 2333 AA; 259645 MW; 4FC667DCC521BC60 CRC64;	
Query Match 4.4%; Score 125.5; DB 1; Length 2333;		
Best Local Similarity 22.1%; Pred.No. 0.53;		
Matches 71; Conservative 34; Mismatches 139; Indels 77; Gaps 10;		
QY	70 NSPGEVLLLELGPENLPYLAHLSRMVNGYAGGVEVQVLLAGNAFTAGKLVFAAAPPHP 129	
Db	358 DKPFGYLTLELPTDHHGVFGHLVDSYATMRNGWDVSVSAVGNQFNGGCLLVAMPWPKA 417	
QY	130 LENISFGQTMPPHVIIDVRLPEVLLPLPDVNRNPFHYNQNEPRMRLVAMLYTFLRSN 189	
Db	418 FDTREKYQLTLFPHPFISPTNTAHTIYVYLGVN--RYDQY----- 457	
QY	190 GSGDDVFTVSCVRLTRPSDFDFNVLVPPTLESKTKPFTLTITIGELTNSRFPIDEL 249	
Db	458 -----KKRKPWTLVWVLSPLTVSVNTAAPQIKV 485	
QY	250 YTPSPNESLVQPQNGRCALDGLQGTTLPLTAICSFRG-----RINKVSGENHVMN 302	
Db	486 YAN-----IAPYVHVA--GELPSKEGIFPVACADGYGLVTTDPKTDAPVYVK--VYN 535	
QY	303 MQVTNINGTFPDPGTGVPAPLGTPTDF-----SKLFGVLVSQRDHNDNACSHDAVIATNSAK 358	
Db	536 PKPTNPGR--FTNLIDVAE--ACPTFLRPDDGKPY--VVTRADDTRLLAKFDVSLAAKMS 591	
QY	359 FTFKLGAIQIGTWEEDDVHN 379	
Db	592 NTLGSGIAQYTYQYSGTINLH 612	

RESULT 15	
FAT2_DROME	
ID FAT2_DROME STANDARD; PRT; 4705 AA.	
AC Q9VW71; Q95S51;	
DT 28-FEB-2003 (Rel. 41, Created)	
DT 28-FEB-2003 (Rel. 41, Last sequence update)	
DT 10-OCT-2003 (Rel. 42, Last annotation update)	
DE Putative fat-like cadherin-related tumor suppressor homolog precursor.	
GN FAT2 OR CG7749.	
OS Drosophila melanogaster (Fruit fly).	
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;	
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;	
OC Ephydroidea; Drosophilidae; Drosophila.	
NCBI TaxID=7227;	
RN [1]	
RP SEQUENCE FROM N.A.	
RC STRAIN=Berkley;	
RX MEDLINE=20196006; PubMed=10731132;	
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,	
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,	
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,	
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,	
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,	
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,	
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,	
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,	
RA Beeson K.Y., Bencos P.V., Berman B.P., Bhandari D., Bolshakov S.,	
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,	
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,	
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,	
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,	
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,	
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,	
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,	
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,	
RA Harris N.L., Harvey D.A., Heiman T.J., Hernandez J.R., Houck J.,	
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,	
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,	
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,	
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,	
RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,	
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,	
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,	
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,	
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,	
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,	
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,	
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,	
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,	
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,	
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,	
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,	
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,	
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;	
RT "The genome sequence of Drosophila melanogaster.";	
RL Science 287:2185-2195 (2000).	
RN [2]	
RP REVISIONS.	
RX MEDLINE=22426069; PubMed=125373752;	
RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,	
RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,	
RA Smith C.D., Tupy J.L., Whitfield B.J., Bayraktaroglu L., Berman B.P.,	
RA Bettencourt B.R., Celniker S.E., de Grey A.D.N.J., Drysdale R.A.,	
RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,	
RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,	
RA Lewis S.E.;	
RT "Annotation of the Drosophila melanogaster euchromatic genome: a	
RT systematic review.";	
RL Genome Biol. 3:RESEARCH0083.1-RESEARCH0083.22 (2002).	
RN [3]	
RP SEQUENCE OF 3837-4705 FROM N.A.	



Db	459	QLDAEKKSSYTLTVSAIDQANVGSRKQSSAKVKISVQDMNDNDPIFENVVKVISINENN-	517
Qy	265	RCALDGELOQTTOQLLPTAICSPGRINQKVSGENHVMNQVTNINGTPDPDTPGVPAPLG	324
Db	518	-----LAGSFVWKLTA-----KDRDSENGSYISYSIANLNAVPE-----	552
Qy	325	TPDFSGKLFGLVLSQRDHDNACRSHDAVIATNS-----AKFTPKLGAIOIGTWEEDDVHI	378
Db	553	IDHFGS-IVKTTSLDFFETMKENYELIIRASDWGLPYRRQTEIKLSIV-----VKDIND	605
Qy	379	NOP-----TKFTPVG	388
Db	606	NRPFERNVCYCKVTKSAPWG	626

Search completed: June 1, 2004, 13:47:21  
Job time : 7.46309 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 31.1485 Seconds

(without alignments)  
5419.275 Million cell updates/sec

Title: US-09-926-799-8

Perfect score: 2843

Sequence: 1 MKMASNDAPSDGAGLVLP.....VNQFYSLAPMTGNGRRRVQ 535

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- SPTREMBL.25.\*
- 1: sp\_archaea.\*
  - 2: sp\_bacteria.\*
  - 3: sp\_fungi.\*
  - 4: sp\_human.\*
  - 5: sp\_invertebrate.\*
  - 6: sp\_mammal.\*
  - 7: sp\_mhc.\*
  - 8: sp\_organelle.\*
  - 9: sp\_phase.\*
  - 10: sp\_plant.\*
  - 11: sp\_rodent.\*
  - 12: sp\_virus.\*
  - 13: sp\_vertebrate.\*
  - 14: sp\_unclassified.\*
  - 15: sp\_rvirus.\*
  - 16: sp\_bacteriap.\*
  - 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2843	100.0	535	12 Q9QMK6	Q9qmk6 chitta viru
2	2820	99.2	535	12 Q911L5	Q911l5 human calic
3	2814	99.0	535	12 Q917X7	Q917x7 norwalk-lik
4	2796	98.3	535	12 Q9V784	Q9v784 norwalk vir
5	2784	97.9	535	12 Q915C6	Q915c6 human calic
6	2758	97.0	535	12 Q91V39	Q91v39 human calic
7	2784	87.3	535	12 Q916B4	Q916e4 human calic
8	2477	87.1	535	12 Q917X4	Q917x4 norwalk-lik
9	2477	87.1	535	12 Q916B6	Q916e6 human calic
10	2476	87.1	535	12 Q917Y3	Q917y3 norwalk-lik
11	2474	87.0	535	12 Q68104	Q68104 hawaii cali
12	2468	86.8	535	12 Q917Y0	Q917y0 norwalk-lik
13	2466	86.7	535	12 Q917Y9	Q917y9 norwalk-lik
14	2452	86.2	535	12 Q91V51	Q91v51 human calic
15	2307.5	81.2	548	12 Q915C7	Q915c7 human calic
16	2300.5	80.9	542	12 Q913B6	Q913b6 human calic

17	2298.5	80.8	542	12 Q66293	Q66293 calicivirus
18	2296.5	80.8	542	12 Q12305	Q12305 snow mounta
19	2296.5	80.8	542	12 Q80RD6	Q80rd6 snow mounta
20	2289.5	80.5	542	12 Q96877	Q96877 snow mounta
21	2258.5	79.4	540	12 Q917X1	Q917x1 norwalk-lik
22	2255.5	79.3	540	12 Q917W8	Q917w8 norwalk-lik
23	2254.5	79.3	538	12 Q911E5	Q911e5 human calic
24	2234.5	78.6	540	12 Q91V50	Q91v50 human calic
25	2123.5	74.7	548	12 Q66296	Q66296 calicivirus
26	2123.5	74.7	548	12 Q8V768	Q8v768 norwalk vir
27	2121.5	74.6	548	12 Q8V0P5	Q8v0p5 human calic
28	2121.5	74.6	548	12 Q88291	Q88291 small round
29	2116.5	74.4	548	12 Q918A1	Q918a1 norwalk-lik
30	2116.5	74.4	548	12 Q8V0P4	Q8v0p4 human calic
31	2111.5	74.3	548	12 Q917Z5	Q917z5 norwalk-lik
32	2110.5	74.2	548	12 Q8V771	Q8v771 norwalk vir
33	2108.5	74.2	548	12 Q917Z2	Q917z2 norwalk-lik
34	2108.5	74.2	548	12 Q8BC87	Q8bc87 human calic
35	2104.5	74.0	548	12 Q83880	Q83880 norwalk vir
36	2104.5	74.0	548	12 Q91V40	Q91v40 human calic
37	2103.5	74.0	548	12 Q916E5	Q916e5 human calic
38	2101.5	73.9	548	12 Q917Z8	Q917z8 norwalk-lik
39	2099.5	73.8	548	12 Q68291	Q68291 human calic
40	2095.5	73.7	548	12 Q80FL1	Q80fl1 human calic
41	2094.5	73.7	548	12 Q8BC90	Q8bc90 human calic
42	2093.5	73.6	548	12 Q918A4	Q918a4 norwalk-lik
43	2092.5	73.6	548	12 Q80FK4	Q80fk4 human calic
44	2090.5	73.5	548	12 Q80FK9	Q80fk9 human calic
45	2090.5	73.5	548	12 Q80FK1	Q80fk1 human calic

#### ALIGNMENTS

#### RESULT 1

Q9QMK6 PRELIMINARY; PRT; 535 AA.

AC Q9QMK6;  
DT 01-MAY-2000 (TremBLrel. 13, Created)  
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Chitta virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=104901;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Aichi76-96;  
RX MEDLINE=20471382; PubMed=11021399;  
RA Kobayashi S., Sakae K., Suzuki Y., Ishiko H., Kamata K., Suzuki K.,  
RA Natori K., Miyamura T., Takeda N.;  
RT "Expression of recombinant capsid proteins of chitta virus, a  
RT genogroup II Norwalk virus, and development of an ELISA to detect the  
RT viral antigen.";  
RL Microbiol. Immunol. 44:687-693 (2000).  
DR EMBL; AB032758; BAA84716.1;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
PFam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 535 AA; 58504 MW; 5CCEBC91BFC7B8B CRC64;

Query Match 100.0%; Score 2843; DB 12; Length 535;

Best Local Similarity 100.0%; Pred. No. 1.1e-220;

Matches 535; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY. 1 MKMASNDAPSDGAGLVLPPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60

Db 1 MKMASNDAPSDGAGLVLPPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60

QY 61 NGEFTVSPRNSPGEVLNLNLGPELNPYLHLSRMYNGYAGGVEVQVLLAGNAFTAGKLV 120

Db 61 NGEFTVSPRNSPGEVLNLNLGPELNPYLHLSRMYNGYAGGVEVQVLLAGNAFTAGKLV 120

121 FAAPVPHPLENISPGQITMFPVHVIDVRLLEPVLPLDVRNNFFHYNQNEPRMLVA 180  
 121 FAAPVPHPLENISPGQITMFPVHVIDVRLLEPVLPLDVRNNFFHYNQNEPRMLVA 180  
 181 MLYTPLRNSGDDVFTVSCRYLTPSPDFDNYLVPPPTLESKTKPFTLPILTIGELTNS 240  
 181 MLYTPLRNSGDDVFTVSCRYLTPSPDFDNYLVPPPTLESKTKPFTLPILTIGELTNS 240  
 241 RFPVPIDELYTSPNESLVVQPNQGRCALDGELOQTTLPTAICSFGRINQKVSNGHV 300  
 241 RFPVPIDELYTSPNESLVVQPNQGRCALDGELOQTTLPTAICSFGRINQKVSNGHV 300  
 301 WNMVNTINGTDFDPTGDPVAPLGTDFSGKLFVLSQRDHNDACRSHDAVIATNSAKFT 360  
 301 WNMVNTINGTDFDPTGDPVAPLGTDFSGKLFVLSQRDHNDACRSHDAVIATNSAKFT 360  
 361 PKLGAIOIGTWBDDVHINQPTKFTPVGLFENEGNQWTLPNYSGALTINMGLAPPVAPT 420  
 361 PKLGAIOIGTWBDDVHINQPTKFTPVGLFENEGNQWTLPNYSGALTINMGLAPPVAPT 420  
 421 FPGQILFRSHITPLKGGVADPVICLLPQEWIQLHYQESAPSDVALIRFTNPDTRGV 480  
 421 FPGQILFRSHITPLKGGVADPVICLLPQEWIQLHYQESAPSDVALIRFTNPDTRGV 480  
 481 LFEAKLHRSYIIVANTGSRPIVVPANGYFRFDTWVNFYSLAPMGNGRRRVQ 535  
 481 LFEAKLHRSYIIVANTGSRPIVVPANGYFRFDTWVNFYSLAPMGNGRRRVQ 535

RESULT 2  
 ID Q91115 PRELIMINARY; PRT; 535 AA.  
 AC Q91115  
 DT 01-DEC-2001 (TremBrel. 19, Created)  
 DT 01-DEC-2001 (TremBrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TremBrel. 25, Last annotation update)  
 DE Capsid protein.  
 OS Human calicivirus NLV/Schwerin 003/00/DE.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OC NCBI\_taxid=168852;  
 RN SEQUENCE FROM N.A.  
 RA Kuenkel U., Schreier E.;  
 RA "Molecular epidemiology of outbreaks of gastroenteritis associated  
 RT with Norwalk-like viruses in Germany."  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF397905; AAK92128.1; --  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 535 AA; 58520 MW; 2E97BBD3F3D5FCC1 CRC64;

Query Match 99.2%; Score 2820; DB 12; Length 535;  
 Best Local Similarity 99.1%; Pred. No. 7.7e-219;  
 Matches 530; Conservative 4; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKWASDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60  
 DB 1 MKWASDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60  
 QY 61 NGEFTVSPRNSGCVLLNLELGPENLNPVLAHLRMYNGVAGVEVQVLLAGNAFTAGKLV 120  
 DB 61 NGEFTVSPRNSGCVLLNLELGPENLNPVLAHLRMYNGVAGVEVQVLLAGNAFTAGKLV 120  
 QY 121 FAAPVPHPLENISPGQITMFPVHVIDVRLLEPVLPLDVRNNFFHYNQNEPRMLVA 180  
 DB 121 FAAPVPHPLENISPGQITMFPVHVIDVRLLEPVLPLDVRNNFFHYNQNEPRMLVA 180  
 181 MLYTPLRNSGDDVFTVSCRYLTPSPDFDNYLVPPPTLESKTKPFTLPILTIGELTNS 240  
 DB 181 MLYTPLRNSGDDVFTVSCRYLTPSPDFDNYLVPPPTLESKTKPFTLPILTIGELTNS 240

QY 241 RFPVPIDELYTSPNESLVVQPNQGRCALDGELOQTTLPTAICSFGRINQKVSNGHV 300  
 DB 241 RFPVPIDELYTSPNESLVVQPNQGRCALDGELOQTTLPTAICSFGRINQKVSNGHV 300  
 QY 301 WNMVNTINGTDFDPTGDPVAPLGTDFSGKLFVLSQRDHNDACRSHDAVIATNSAKFT 360  
 DB 301 WNMVNTINGTDFDPTGDPVAPLGTDFSGKLFVLSQRDHNDACRSHDAVIATNSAKFT 360  
 QY 361 PKLGAIOIGTWBDDVHINQPTKFTPVGLFENEGNQWTLPNYSGALTINMGLAPPVAPT 420  
 DB 361 PKLGAIOIGTWBDDVHINQPTKFTPVGLFENEGNQWTLPNYSGALTINMGLAPPVAPT 420  
 QY 421 FPGQILFRSHITPLKGGVADPVICLLPQEWIQLHYQESAPSDVALIRFTNPDTRGV 480  
 DB 421 FPGQILFRSHITPLKGGVADPVICLLPQEWIQLHYQESAPSDVALIRFTNPDTRGV 480  
 QY 481 LFEAKLHRSYIIVANTGSRPIVVPANGYFRFDTWVNFYSLAPMGNGRRRVQ 535  
 DB 481 LFEAKLHRSYIIVANTGSRPIVVPANGYFRFDTWVNFYSLAPMGNGRRRVQ 535

RESULT 3  
 ID Q917X7 PRELIMINARY; PRT; 535 AA.  
 AC Q917X7  
 DT 01-DEC-2001 (TremBrel. 19, Created)  
 DT 01-DEC-2001 (TremBrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TremBrel. 25, Last annotation update)  
 DE Capsid protein.  
 GN ORF2.  
 OS Norwalk-like virus NLV/Honolulu/314/1994/US.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OC NCBI\_taxid=171852;  
 RN SEQUENCE FROM N.A.  
 RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
 RA MEDLINE=97193806; PubMed=9041391;  
 RA "A one-tube method of reverse transcription-PCR to efficiently amplify  
 RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
 RT of small round-structured viruses (Norwalk-like viruses).";  
 RL J. Clin. Microbiol. 35:570-577(1997).  
 RN [2]  
 RP SEQUENCE OF 100-192 FROM N.A.  
 RC STRAIN=Hu/NLV/Honolulu/314/1994/US;  
 RX MEDLINE=98071277; PubMed=9407386;  
 RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,  
 RA Seto Y., Monroe S.S., Glass R.I.;  
 RT "Correlation of patient immune responses with genetically  
 RT characterized small round-structured viruses involved in outbreaks of  
 RT nonbacterial acute gastroenteritis in the United States, 1990 to  
 RT 1995.";  
 RL J. Med. Virol. 53:372-383(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/Honolulu/314/1994/US;  
 RX MEDLINE=20266071; PubMed=10804147;  
 RA Ando T., Noel J.S., Fankhauser R.L.;  
 RT "Genetic classification of 'Norwalk'-like viruses.";  
 RL J. Infect. Dis. 181:S336-S348(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/Honolulu/314/1994/US;  
 RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;  
 RA Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF414420; AAL13013.1; --  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 535 AA; 58519 MW; C3B7A70CF3617F71 CRC64;

```
Query Match      99.0%; Score 2814; DB 12; Length 535;
Best Local Similarity 98.9%; Pred. No. 2.3e-218;
Matches 529; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60

QY 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVYAGGVEVQVLLAGNAFTAGKLV 120
DB 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVYAGGVEVQVLLAGNAFTAGKLV 120

QY 121 FFAVPPHPLENIISQGITMFPVHVIDVRLTLEPVLLPDPVNNFFHYNQNEPRMRLVA 180
DB 121 FFAVPPHPLENIISQGITMFPVHVIDVRLTLEPVLLPDPVNNFFHYNQNEPRMRLVA 180

QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTKPFTLPILTIGELTNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTKPFTLPILTIGELTNS 240

QY 241 RPPVPIDELITYSPNESLVVQNGRCALDGELOGTTLPTAICSFGRINQKVGSENVH 300
DB 241 RPPVPIDELITYSPNESLVVQNGRCALDGELOGTTLPTAICSFGRINQKVGSENVH 300

QY 301 WNMQVTNINGTTPDPTEDVPAPLGTDFSGKLVGLSQRDHDNACRSHDAVIATNSAKFT 360
DB 301 WNMQVTNINGTTPDPTEDVPAPLGTDFSGKLVGLSQRDHDNACRSHDAVIATNSAKFT 360

QY 361 PKLGAIGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTINMGLAPPVAPT 420
DB 361 PKLGAIGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTINMGLAPPVAPT 420

QY 421 PFGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSDVALIRFTNPDTCGRV 480
DB 421 PFGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSDVALIRFTNPDTCGRV 480

QY 481 LFEAKLHRSGYITVANTGSRPIVWPANGYFRFDTWVNOFYSLAPMGTGNGRRRVQ 535
DB 481 LFEAKLHRSGYITVANTGSRPIVWPANGYFRFDTWVNOFYSLAPMGTGNGRRRVQ 535

RESULT 4
Q8V784 PRELIMINARY; PRT; 535 AA.
AC Q8V784;
DT 01-MAR-2002 (TremBLrel. 20, Created)
DT 01-MAR-2002 (TremBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE ORF2 protein (Capsid).
OS Norwalk virus, and
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983, 95340;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk virus; STRAIN=UIGII;
RA Kojima S., Kageyama T., Fukushima S., Hoshino F.B., Shinohara M.,
RA Uchida K., Natori K., Takeda N., Katayama K.;
RT "Genogroup-specific primers for detect Norwalk like virus.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DDBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama UI;
RA Katayama K.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama UI;
RX MEDLINE=22192455; PubMed=12202225;
RA Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA Hoshino F.B., Fukushima S., Shinohara M., Uchida K., Suzuki Y.,
RA Gojobori T., Takeda N.;
```

```
"Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
Viruses.";
RT Virology 299:225-239 (2002).
DR EMBL; AB067536; BAB84137.1; -.
DR EMBL; AB039775; BAC11816.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58615 MW; 6B0B2D3162436787 CRC64;

Query Match      98.3%; Score 2796; DB 12; Length 535;
Best Local Similarity 98.3%; Pred. No. 6.7e-217;
Matches 526; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60

QY 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVYAGGVEVQVLLAGNAFTAGKLV 120
DB 61 NGEFTVSPRNSPGEVLLNLELGPENLPYLAHLSRMVYAGGVEVQVLLAGNAFTAGKLV 120

QY 121 FFAVPPHPLENIISQGITMFPVHVIDVRLTLEPVLLPDPVNNFFHYNQNEPRMRLVA 180
DB 121 FFAVPPHPLENIISQGITMFPVHVIDVRLTLEPVLLPDPVNNFFHYNQNEPRMRLVA 180

QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTKPFTLPILTIGELTNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTKPFTLPILTIGELTNS 240

QY 241 RPPVPIDELITYSPNESLVVQNGRCALDGELOGTTLPTAICSFGRINQKVGSENVH 300
DB 241 RPPVPIDELITYSPNESLVVQNGRCALDGELOGTTLPTAICSFGRINQKVGSENVH 300

QY 301 WNMQVTNINGTTPDPTEDVPAPLGTDFSGKLVGLSQRDHDNACRSHDAVIATNSAKFT 360
DB 301 WNMQVTNINGTTPDPTEDVPAPLGTDFSGKLVGLSQRDHDNACRSHDAVIATNSAKFT 360

QY 361 PKLGAIGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTINMGLAPPVAPT 420
DB 361 PKLGAIGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTINMGLAPPVAPT 420

QY 421 PFGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSDVALIRFTNPDTCGRV 480
DB 421 PFGEQILFFRSHIPLKGGVADPVIDCLLPQEWIQLHYQESAPSDVALIRFTNPDTCGRV 480

QY 481 LFEAKLHRSGYITVANTGSRPIVWPANGYFRFDTWVNOFYSLAPMGTGNGRRRVQ 535
DB 481 LFEAKLHRSGYITVANTGSRPIVWPANGYFRFDTWVNOFYSLAPMGTGNGRRRVQ 535

RESULT 5
Q915C6 PRELIMINARY; PRT; 535 AA.
ID Q915C6
AC Q915C6;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Pirna/110/00/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=173930;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Pirna/110/00/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RT with Norwalk-like viruses in Germany.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AF427119; AAL18875.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
```

```
DR Pfam; PF00915; Calici_coat; 1. D039C359F86DC1F9 CRC64;
SQ SEQUENCE 535 AA; 58505 MW; 2784; DB 12; Length 535;

Query Match
Best Local Similarity 97.4%; Score 2784; DB 12; Length 535;
Matches 521; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
QY 61 NGEFTVSPNSGCVLLNLLELGPENLPYLAHLRMYNGYAGGVEQVLLAGNAFTAGKLV 120
Db 61 NGEFTVSPNSGCVLLNLLELGPENLPYLAHLRMYNGYAGGVEQVLLAGNAFTAGKLV 120
QY 121 FFAVPPHFFLENISQGITMFFPHVIVDVTLEPVLPLPDVRRNPFHYNQNEPRMRLVA 180
Db 121 FFAVPPHFFLENISQGITMFFPHVIVDVTLEPVLPLPDVRRNPFHYNQNEPRMRLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTKPFTLPILITIGELTNS 240
Db 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTKPFTLPILITIGELTNS 240
QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTKPFTLPILITIGELTNS 240
Db 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTKPFTLPILITIGELTNS 240
QY 241 RFPVPIDELYTSNESLVVQPNQRCALDGELOQTTLPTAICSFGRINQKVS GENHV 300
Db 241 RFPVPIDELYTSNESLVVQPNQRCALDGELOQTTLPTAICSFGRINQKVS GENHV 300
QY 241 RFPVPIDELYTSNESLVVQPNQRCALDGELOQTTLPTAICSFGRINQKVS GENHV 300
Db 241 RFPVPIDELYTSNESLVVQPNQRCALDGELOQTTLPTAICSFGRINQKVS GENHV 300
QY 301 WNNQVTNINGTPDPTGDPVAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
Db 301 WNNQVTNINGTPDPTGDPVAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
QY 301 WNNQVTNINGTPDPTGDPVAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
Db 301 WNNQVTNINGTPDPTGDPVAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
QY 361 PKLGAIGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTLNMG LAPVAPT 420
Db 361 PKLGAIGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTLNMG LAPVAPT 420
QY 421 FPGEQILFFRSHIPLKGVADPVIDCLLPQEWIQLYQESAPSQSDVALIRFNTPTGRV 480
Db 421 FPGEQILFFRSHIPLKGVADPVIDCLLPQEWIQLYQESAPSQSDVALIRFNTPTGRV 480
QY 481 LFEAKLHRSYITVANTGSRPIVVPANGYFRFDTWVNOFYSLAPMG TGNGRRRVQ 535
Db 481 LFEAKLHRSYITVANTGSRPIVVPANGYFRFDSVWVNOFYSLAPMG TGNGRRRVQ 535

RESULT 6
Q91V39 PRELIMINARY; PRT; 535 AA.
ID Q91V39
AC Q91V39;
DT 01-OCT-2000 (TremBLrel. 15, Created)
DT 01-OCT-2000 (TremBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus HU/NLV/Wortley/90/UK.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=122922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HU/NLV/Wortley/90/UK;
RX MEDLINE=20404883; PubMed=10949950;
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,
RA Clegg J.C., Chamberlain J., Brown D.W.G.;
RT "Capsid protein diversity among 'Norwalk-like' viruses.";
RL Virus Genes 20:227-236(2000).
DR EMBL; AJ277618; CAB89099.1; -
DR GO; GO:004601; F:peroxidase activity; IEA.
DR GO; GO:006979; P:response to oxidative stress; IEA.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR020206; Peroxidase.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR PROSITE; PS00435; PEROXIDASE 1; 1.
SQ SEQUENCE 535 AA; 58487 MW; 29B513DFD2EDCFD2 CRC64;
```

```
Query Match
Best Local Similarity 97.0%; Score 2758; DB 12; Length 535;
Matches 518; Conservative 8; Mismatches 9; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
QY 61 NGEFTVSPNSGCVLLNLLELGPENLPYLAHLRMYNGYAGGVEQVLLAGNAFTAGKLV 120
Db 61 NGEFTVSPNSGCVLLNLLELGPENLPYLAHLRMYNGYAGGVEQVLLAGNAFTAGKLV 120
QY 121 FFAVPPHFFLENISQGITMFFPHVIVDVTLEPVLPLPDVRRNPFHYNQNEPRMRLVA 180
Db 121 FFAVPPHFFLENISQGITMFFPHVIVDVTLEPVLPLPDVRRNPFHYNQNEPRMRLVA 180
QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTKPFTLPILITIGELTNS 240
Db 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNVLPVPTLESKTKPFTLPILITIGELTNS 240
QY 241 RFPVPIDELYTSNESLVVQPNQRCALDGELOQTTLPTAICSFGRINQKVS GENHV 300
Db 241 RFPVPIDELYTSNESLVVQPNQRCALDGELOQTTLPTAICSFGRINQKVS GENHV 300
QY 301 WNNQVTNINGTPDPTGDPVAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
Db 301 WNNQVTNINGTPDPTGDPVAPLGTDFSGKLFVLSQRDHDNACRSHDAVIATNSAKFT 360
QY 361 PKLGAIGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTLNMG LAPVAPT 420
Db 361 PKLGAIGTWEEDDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTLNMG LAPVAPT 420
QY 421 FPGEQILFFRSHIPLKGVADPVIDCLLPQEWIQLYQESAPSQSDVALIRFNTPTGRV 480
Db 421 FPGEQILFFRSHIPLKGVADPVIDCLLPQEWIQLYQESAPSQSDVALIRFNTPTGRV 480
QY 481 LFEAKLHRSYITVANTGSRPIVVPANGYFRFDTWVNOFYSLAPMG TGNGRRRVQ 535
Db 481 LFEAKLHRSYITVANTGSRPIVVPANGYFRFDSVWVNOFYSLAPMG TGNGRRRVQ 535

RESULT 7
Q916E4 PRELIMINARY; PRT; 535 AA.
ID Q916E4
AC Q916E4;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Wiesbaden 294/01/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=173922;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NLV/Wiesbaden 294/01/DE;
RA Kuenkel U., Hoehne M., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated
RT with Norwalk-like viruses in Germany.";
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP425769; AAL18862.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58795 MW; 987237124A0605B2 CRC64;

Query Match
Best Local Similarity 85.0%; Score 2481; DB 12; Length 535;
Matches 455; Conservative 37; Mismatches 43; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
```

Db 1 MKMASNDAAPSNDGAAGLVPEVNNETMALEVPVAGASIAAPLTGQNNVIDPWRMNFVQAP 60  
 QY 61 NGEFTVSPRNSGEVLLNLELGPENLPYLAHLSRMVYAGGVEVQVLLAGNAFTAGKLV 120  
 Db 61 NGEFTVSPRNSGEVLLNLELGPENLPYLAHLSRMVYAGGVEVQVLLAGNAFTAGKLV 120  
 QY 121 FAAPVPPHLENISPGQITMPPHVIDVRLTLEPVLPLPVDVNNFFHYNQNEPRMRLVA 180  
 Db 121 FAAPVPPHLENISPGQITMPPHVIDVRLTLEPVLPLPVDVNNFFHYNQNEPRMRLVA 180  
 QY 181 MLYTLRNSGGDDVFTVSCRVLTRPSDDFDFNVLVPPVLESKTKPFTLPILTIGELTNS 240  
 Db 181 MLYTLRNSGGDDVFTVSCRVLTRPSDDFDFNVLVPPVLESKTKPFTLPILTIGELTNS 240  
 QY 241 RFPVPEIDELTSPNESLVVQPNQRCALDGLGQTTQLLTAICSFGRINQKVS GENHV 300  
 Db 241 RFPVPEIDELTSPNESLVVQPNQRCALDGLGQTTQLLTAICSFGRINQKVS GENHV 300  
 QY 301 WNMQVNTINGPFPDTPGVAPLGPDPFSGKLVGVSORDHNDACRSHDAVIATNSAKPT 360  
 Db 301 WNMQVNTINGPFPDTPGVAPLGPDPFSGKLVGVSORDHNDACRSHDAVIATNSAKPT 360  
 QY 361 PKLGAIGTWEEDDVHINOPTKTPVGLFENEGENOWTLPNYSICALTNMGLAPPVAPT 420  
 Db 361 PKLGAIGTWEEDDVHINOPTKTPVGLFENEGENOWTLPNYSICALTNMGLAPPVAPT 420  
 QY 421 FPGQILFFRSHIPLKGGTSNGAIDCLLPQEWVQHFYQESAPSSDVALIRYTNPDTCGRV 480  
 Db 421 FPGQILFFRSHIPLKGGTSNGAIDCLLPQEWVQHFYQESAPSSDVALIRYTNPDTCGRV 480  
 QY 481 LFEAKLHRSYGITVANTGSRPIVVPANGYFRPDTWVNOFYSLAPMGTCNGRRRVQ 535  
 Db 481 LFEAKLHRSYGITVANTGSRPIVVPANGYFRPDTWVNOFYSLAPMGTCNGRRRVQ 535

RESULT 8  
 Q917X4  
 ID Q917X4 PRELIMINARY; PRT; 535 AA.  
 AC Q917X4  
 DT 01-DEC-2001 (TREMELrel. 19, Created)  
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Capsid protein.  
 GN ORF2.  
 OS Norwalk-like virus NLV/Port Canaveral/301/1994/US.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OC NCBI\_TaxID=171853;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;  
 RX MEDLINE=97193806; PubMed=9041391;  
 RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
 a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
 of small round-structured viruses (Norwalk-like viruses).";  
 RL J. Clin. Microbiol. 35:570-577(1997).  
 RN [2]  
 RP SEQUENCE OF 100-192 FROM N.A.  
 RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;  
 RX MEDLINE=98071277; PubMed=9407386;  
 RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,  
 RA Seto Y., Monroe S.S., Glass R.I.;  
 RT "Correlation of patient immune responses with genetically  
 characterized small round-structured viruses involved in outbreaks of  
 nonbacterial acute gastroenteritis in the United States, 1990 to  
 1995.";  
 RL J. Med. Virol. 53:372-383(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;  
 RX MEDLINE=20266071; PubMed=10804147;  
 RA Ando T., Noel J.S., Fankhauser R.L.;

"Genetic classification of 'Norwalk-like viruses.';  
 J. Infect. Dis. 181:S336-S348(2000).  
 [4]

SEQUENCE FROM N.A.  
 STRAIN=Hu/NLV/Port Canaveral/301/1994/US;  
 RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glaes R.I., Fankhauser R.L.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF414421; AAL13016.1; -;  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 535 AA; 58738 MW; 7C29B02DFDC724C6 CRC64;

Query Match 87.1%; Score 2477; DB 12; Length 535;  
 Best Local Similarity 85.0%; Pred. No. 3.6e-191;  
 Matches 455; Conservative 36; Mismatches 44; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAAGLVPEVNNETMALEVPVAGASIAAPLTGQNNVIDPWRMNFVQAP 60  
 Db 1 MKMASNDAAPSNDGAAGLVPEVNNETMALEVPVAGASIAAPLTGQNNVIDPWRMNFVQAP 60  
 QY 61 NGEFTVSPRNSGEVLLNLELGPENLPYLAHLSRMVYAGGVEVQVLLAGNAFTAGKLV 120  
 Db 61 NGEFTVSPRNSGEVLLNLELGPENLPYLAHLSRMVYAGGVEVQVLLAGNAFTAGKLV 120  
 QY 121 FAAPVPPHLENISPGQITMPPHVIDVRLTLEPVLPLPVDVNNFFHYNQNEPRMRLVA 180  
 Db 121 FAAPVPPHLENISPGQITMPPHVIDVRLTLEPVLPLPVDVNNFFHYNQNEPRMRLVA 180  
 QY 181 MLYTLRNSGGDDVFTVSCRVLTRPSDDFDFNVLVPPVLESKTKPFTLPILTIGELTNS 240  
 Db 181 MLYTLRNSGGDDVFTVSCRVLTRPSDDFDFNVLVPPVLESKTKPFTLPILTIGELTNS 240  
 QY 241 RFPVPEIDELTSPNESLVVQPNQRCALDGLGQTTQLLTAICSFGRINQKVS GENHV 300  
 Db 241 RFPVPEIDELTSPNESLVVQPNQRCALDGLGQTTQLLTAICSFGRINQKVS GENHV 300  
 QY 301 WNMQVNTINGPFPDTPGVAPLGPDPFSGKLVGVSORDHNDACRSHDAVIATNSAKPT 360  
 Db 301 WNMQVNTINGPFPDTPGVAPLGPDPFSGKLVGVSORDHNDACRSHDAVIATNSAKPT 360  
 QY 361 PKLGAIGTWEEDDVHINOPTKTPVGLFENEGENOWTLPNYSICALTNMGLAPPVAPT 420  
 Db 361 PKLGAIGTWEEDDVHINOPTKTPVGLFENEGENOWTLPNYSICALTNMGLAPPVAPT 420  
 QY 421 FPGQILFFRSHIPLKGGTSNGAIDCLLPQEWVQHFYQESAPSSDVALIRYTNPDTCGRV 480  
 Db 421 FPGQILFFRSHIPLKGGTSNGAIDCLLPQEWVQHFYQESAPSSDVALIRYTNPDTCGRV 480  
 QY 481 LFEAKLHRSYGITVANTGSRPIVVPANGYFRPDTWVNOFYSLAPMGTCNGRRRVQ 535  
 Db 481 LFEAKLHRSYGITVANTGSRPIVVPANGYFRPDTWVNOFYSLAPMGTCNGRRRVQ 535

RESULT 9

Q916E6  
 ID Q916E6 PRELIMINARY; PRT; 535 AA.  
 AC Q916E6  
 DT 01-DEC-2001 (TREMELrel. 19, Created)  
 DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
 DE Capsid protein.  
 OS Human calicivirus NLV/Dillingen 391/01/DE.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OC NCBI\_TaxID=173920;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NLV/Dillingen 391/01/DE;  
 RA Kuenkel U., Hoehne M., Schreier E.;  
 RT "Molecular epidemiology of outbreaks of gastroenteritis associated  
 with Norwalk-like viruses in Germany.";  
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.

RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Bates M.K.,  
 RA Seto Y., Monroe S.S., Glass R.I.,  
 RT "Correlation of patient immune responses with genetically  
 RT characterized small round-structured viruses involved in outbreaks of  
 RT nonbacterial acute gastroenteritis in the United States, 1990 to  
 RT 1995.";  
 RL J. Med. Virol. 53:372-383 (1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/Westover/302/1994/US;  
 RX MEDLINE=20266071; PubMed=10804147;  
 RA Ando T., Noel J.S., Fankhauser R.L.;  
 RT "Genetic classification of 'Norwalk'-like viruses.";  
 RL J. Infect. Dis. 181:S336-S348(2000).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/Westover/302/1994/US;  
 RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF414418; AAL13007.1; --  
 DR InterPro: IPR004005; Calici\_coat.  
 DR InterPro: IPR008975; Viral\_Cap\_coat.  
 DR Pfam: PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 535 AA; 58685 MW; 62620468ABB44EC0 CRC64;

Query Match 87.1%; Score 2477; DB 12; Length 535;  
 Best Local Similarity 84.9%; Pred. No. 3.6e-191;  
 Matches 454; Conservative 37; Mismatches 44; Indels 0; Gaps 0;  
 QY 1 MKMASNDAAPSNDGAAGLVPEANNTMALEPVAGASIAAAPTGTQNNIIDPWIRLNFVQAP 60  
 Db 1 MKMASNDAAPSNDGAAGLVPEANNTMALEPVAGASIAAAPTGTQNNIIDPWIRLNFVQAP 60  
 QY 61 NGEFTVSPNSGCEVLLNLELGPENLPYLAHLSRMVNGYAGGVEVQVLLAGNAFTAGKLV 120  
 Db 61 NGEFTVSPNSGCEVLLNLELGPENLPYLAHLSRMVNGYAGGVEVQVLLAGNAFTAGKLV 120  
 QY 121 FAAPVPHPELENISPGQITMFPFVHVIDVRLTPVLLPLPDVVRNFFHYNQNEPRMRLVA 180  
 Db 121 FAAPVPHPELENISPGQITMFPFVHVIDVRLTPVLLPLPDVVRNFFHYNQNEPRMRLVA 180  
 QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDNLYVPPTLESKTKPFTLPILITIGELTNS 240  
 Db 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDNLYVPPTLESKTKPFTLPILITIGELTNS 240  
 QY 241 RFPVPIDELTSPNESLVVQPNQRCALDGELOGTTLPTAICSPGRINOKVSGENHV 300  
 Db 241 RFPVPIDELTSPNESLVVQPNQRCALDGELOGTTLPTAICSPGRINOKVSGENHV 300  
 QY 301 WNMQVNTINGTPDPTGDPAPLGTDPFSGKLVSRQDHNA CRSHDAVIATNSAKFT 360  
 Db 301 WNMQVNTINGTPDPTGDPAPLGTDPFSGKLVSRQDHNA CRSHDAVIATNSAKFT 360  
 QY 361 PKLGAIOIGTWEEEDVHINQPTKFTPVGLFENEGFNQWTLPNYSGALTLMGLAPPVAPT 420  
 Db 361 PKLGSVVLGTWEDRDFINQPTKFTPVGLYDTHFNQWALPNYSGALTLMGLAPPVAPT 420  
 QY 421 FPGQILFFRSHIPLKGGVADPVIDCLLPQEWIOHLYQESAPSDVALIRFTNPDTRGV 480  
 Db 421 FPGQILFFRSHIPLKGGTNGAIDCLLPQEWVQHFYQESAPSDVALIRFTNPDTRGV 480  
 QY 481 LFEAKLHRSGYITVANTGSRPIVVPNGYFRFDSVWVNFYSLAPMGTCNGRRVQ 535  
 Db 481 LFEAKLHRQGITVANGSRPIVVPNGYFRFDSVWVNFYSLAPMGTCNGRRVQ 535

RESULT 10  
 Q917Y3 ID Q917Y3 PRELIMINARY; PRT; 535 AA.  
 AC Q917Y3;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Capsid protein.  
 GN ORF2.  
 OS Norwalk-like virus NLV/Westover/302/1994/US.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OC NCBI\_TaxID=171850;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/Westover/302/1994/US;  
 RX MEDLINE=97193806; PubMed=9041391;  
 RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
 RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
 RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
 RT of small round-structured viruses (Norwalk-like viruses).";  
 RL J. Clin. Microbiol. 35:570-577 (1997).  
 RN [2]  
 RP SEQUENCE OF 100-192 FROM N.A.  
 RC STRAIN=Hu/NLV/Westover/302/1994/US;  
 RX MEDLINE=98071277; PubMed=9407386;

RESULT 11  
 Q68104 ID Q68104 PRELIMINARY; PRT; 535 AA.  
 AC Q68104;  
 DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Capsid protein.  
OS Hawaii calicivirus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=33750;  
RN [1]\_TaxID=33750;  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;  
RX MEDLINE=94358485; PubMed=8077710;  
RA Lew J.F., Kapikian A.Z., Valdesuso J., Green K.Y.;  
RT "Molecular characterization of Hawaii virus and other Norwalk-like  
RT viruses; evidence for genetic polymorphism among human  
RT caliciviruses."; J. Infect. Dis. 170:535-542 (1994).  
RL J. Infect. Dis. 170:535-542 (1994).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;  
RA Lew J.F.;  
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;  
RA Green K.Y., Sosnovtsev S.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;  
RA Pleitneva M.A., Sosnovtsev S., Green K.Y.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; U07611; AAB97768.2; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 535 AA; 58741 MW; D0CA252A210E4599 CRC64;

Query Match 87.0%; Score 2474; DB 12; Length 535;  
Best Local Similarity 84.9%; Pred. No. 6.3e-191;  
Matches 454; Conservative 40; Mismatches 41; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIDPWIRLNFVQAP 60  
DB 1 MKMASNDAAPSNDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIDPWIRLNFVQAP 60  
QY 61 NGEFTVSPNSPGEVLLNLELGPENLPYLAHLRMYNGYAGGVEVQVLLAGNAFTAGKLV 120  
DB 61 NGEFTVSPNSPGEVLLNLELGPENLPYLAHLRMYNGYAGGVEVQVLLAGNAFTAGKLV 120  
QY 121 FAAPVPHPPLENISPGQITMFPFHVIDVRLTLEPVLPLPDVRNFFHYNQOEPRMRLVA 180  
DB 121 FAAPVPHPPLENISPGQITMFPFHVIDVRLTLEPVLPLPDVRNFFHYNQOEPRMRLVA 180  
QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDNYLVPPTLESKTKPFTLPILITIGELTNS 240  
DB 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDNYLVPPTLESKTKPFTLPILITIGELTNS 240  
QY 241 RFPVPIDELTSPNSLSVQPNQRCALDGEQGTQLLTALTCSPGRINOKVSGENHV 300  
DB 241 RFPVPIDELTSPNSGVIVQPNQGRSTLDGELLGTTQLVPSNICALGRINAQVDDHQ 300  
QY 301 WNVQVNTINGTDPDTPGVAPLTPDPSGKLVGVLSQRDHNAACRSHDAVATNSAKFT 360  
DB 301 WNVQVNTINGTDPDTPGVAPLTPDPSGKLVGVLSQRDHNAACRSHDAVATNSAKFT 360  
QY 361 PKLGAIOGTWEEDVHINQPKFTPVGLFENEGFNQWTLPNYSGALTLNMGGLAPVAPT 420  
DB 361 PKLGSVILGTWEESDLNQNPTFRFTVGLFNTDHDQWALPVSGLRTLNMMNLVSPL 420  
QY 421 FPGEQILFRSHIPLKGGVADPVDCLLPQEWIQHLYQESAPSQSQVALIRTNPDTRGV 480  
DB 421 FPGEQILFRSHIPLKGGTSDCAIDCLLPQEWIQHLYQESAPSPTDVALIRTNPDTRGV 480  
QY 481 LFEAKLHRSGYITVANTGSRPIVWPANGYFRFDTWVNFYSLAPMGTCNGRRRVQ 535  
DB 481 LFEAKLHRSGYITVANTGSRPIVWPANGYFRFDTWVNFYSLAPMGTCNGRRRVQ 535

Db 481 LFEAKLHRSQFTTVANSGRPIVVPNGYFRPDSWVNFYSLAPMGTCNGRRRVQ 535

RESULT 12  
Q917Y0 PRELIMINARY; PRT; 535 AA.  
AC Q917Y0;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
DE Capsid protein.  
GN ORF2.  
OS Norwalk-like virus NLV/Richmond/283/1994/US.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=171851;  
RN [1]\_TaxID=171851;  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Richmond/283/1994/US;  
RX MEDLINE=97193806; PubMed=9041391;  
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
RT of small round-structured viruses (Norwalk-like viruses).";  
RL J. Clin. Microbiol. 35:570-577 (1997).  
RN [2]  
RP SEQUENCE OF 100-192 FROM N.A.  
RC STRAIN=Hu/NLV/Richmond/283/1994/US;  
RX MEDLINE=98071277; PubMed=9407386;  
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,  
RA Seto Y., Monroe S.S., Glass R.I.;  
RT "Correlation of patient immune responses with genetically  
RT characterized small round-structured viruses involved in outbreaks of  
RT nonbacterial acute gastroenteritis in the United States, 1990 to  
RT 1995."; J. Med. Virol. 53:372-383 (1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Richmond/283/1994/US;  
RX MEDLINE=20266071; PubMed=10804147;  
RA Ando T., Noel J.S., Fankhauser R.L.;  
RT "Genetic classification of 'Norwalk'-like viruses.";  
RL J. Infect. Dis. 181:S336-S348 (2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Richmond/283/1994/US;  
RA Ando T., Seto Y., Noel J.S., Glass R.I., Fankhauser R.L.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF414419; AAL13010.1; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 535 AA; 58742 MW; 4EDACABCB8E22F80 CRC64;

Query Match 86.8%; Score 2468; DB 12; Length 535;  
Best Local Similarity 84.9%; Pred. No. 1.9e-190;  
Matches 454; Conservative 36; Mismatches 45; Indels 0; Gaps 0;

QY 1 MKMASNDAAPSNDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIDPWIRLNFVQAP 60  
DB 1 MKMASNDAAPSNDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIDPWIRLNFVQAP 60  
QY 61 NGEFTVSPNSPGEVLLNLELGPENLPYLAHLRMYNGYAGGVEVQVLLAGNAFTAGKLV 120  
DB 61 NGEFTVSPNSPGEVLLNLELGPENLPYLAHLRMYNGYAGGVEVQVLLAGNAFTAGKLV 120  
QY 121 FAAPVPHPPLENISPGQITMFPFHVIDVRLTLEPVLPLPDVRNFFHYNQOEPRMRLVA 180  
DB 121 FAAPVPHPPLENISPGQITMFPFHVIDVRLTLEPVLPLPDVRNFFHYNQOEPRMRLVA 180  
QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDNYLVPPTLESKTKPFTLPILITIGELTNS 240  
DB 181 MLYTPLRSNGSGDDVFTVSCRVLTRPSDFDNYLVPPTLESKTKPFTLPILITIGELTNS 240

QY 241 RFPVPIDELYTSNLSVQNGRCALDGELOCTTOLPTAICSFRGRINOKVSGENHV 300  
DB 241 RFPVPIDELYTSNLSVQNGRCALDGELOCTTOLPTAICSFRGRINOKVSGENHV 300  
QY 301 MNQVNTNNGTFFDPTGVPAPLGPDPFSGKLFVLSQRDHDNACRSHDAVIANSKFT 360  
DB 301 MNQVNTNNGTFFDPTGVPAPLGPDPFSGKLFVLSQRDHDNACRSHDAVIANSKFT 360  
QY 361 PKLGAIOIGTWEDDDVHINQPTKFTPVGLFENEGFNQWTLNYSGALTNNGLAPPVAPT 420  
DB 361 PKLGSVVLGTWEDRDFINQPARFPTVGLYTDHFNQWALPNYSGALTNNGLAPPVAPT 420  
QY 421 FPGEQILFRSHIPLKGGVADPVIDCLLPQEWIQHLYQESAPSDQDVALIRFTNPDTCRV 480  
DB 421 FPGEQILFRSHIPLKGGVADPVIDCLLPQEWIQHLYQESAPSDQDVALIRFTNPDTCRV 480  
QY 481 LFEAKLHRSQGIITVANTGSRPIVVPANGYFRDFTWVWQFYSLAPMGNGRRVQ 535  
DB 481 LFEAKLHRSQGIITVANTGSRPIVVPANGYFRDFTWVWQFYSLAPMGNGRRVQ 535  
RESULT 13  
Q917Y9 PRELIMINARY; PRT; 535 AA.  
ID Q917Y9  
AC Q917Y9  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
GN ORF2.  
OS Norwalk-like virus NLV/Miami/81/1986/US.  
OC Viruses; serNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OC NCBI\_taxid=171848;  
RX [1]  
RC SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Miami/81/1986/US;  
RX MEDLINE=97193806; PubMed=9041391;  
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
RT of small round-structured viruses (Norwalk-like viruses).";  
RL J. Clin. Microbiol. 35:570-577(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Miami/81/1986/US;  
RX MEDLINE=20266071; PubMed=10804147;  
RA Ando T., Noel J.S., Fankhauser R.L.;  
RT "Genetic classification of 'Norwalk-like viruses.'";  
RL J. Infect. Dis. 181:S336-S348(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Miami/81/1986/US;  
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Fankhauser R.L.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF414416; AAL13001.1;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 535 AA; 58745 MW; D43CDF7BD8C7225A CRC64;  
Query Match 86.7%; Score 2466; DB 12; Length 535;  
Best Local Similarity 84.7%; Pred. No. 2.8e-190;  
Matches 453; Conservative 37; Mismatches 45; Indels 0; Gaps 0;  
QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60  
DB 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60  
QY 61 NGEFTVSPRNSPQEVILNLELGPENLPLAHLRSMYNGYAGGVQVLLAGNAFTAGKLV 120  
DB 61 NGEFTVSPRNSPQEVILNLELGPENLPLAHLRSMYNGYAGGVQVLLAGNAFTAGKLV 120

QY 121 FAAPDPHPLENI SFGQITMFPVHIIDVRLTEPVLPLPDVRNFFHYNQNEPRMLVA 180  
DB 121 FAAPDPHPLENI SFGQITMFPVHIIDVRLTEPVLPLPDVRNFFHYNQNEPRMLVA 180  
QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNFLVLPETVESKTPFTLPILTIGELTNS 240  
DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNFLVLPETVESKTPFTLPILTIGELTNS 240  
QY 241 RFPVPIDELYTSNLSVQNGRCALDGELOCTTOLPTAICSFRGRINOKVSGENHV 300  
DB 241 RFPVPIDELYTSNLSVQNGRCALDGELOCTTOLPTAICSFRGRINOKVSGENHV 300  
QY 301 MNQVNTNNGTFFDPTGVPAPLGPDPFSGKLFVLSQRDHDNACRSHDAVIANSKFT 360  
DB 301 MNQVNTNNGTFFDPTGVPAPLGPDPFSGKLFVLSQRDHDNACRSHDAVIANSKFT 360  
QY 361 PKLGAIOIGTWEDDDVHINQPTKFTPVGLFENEGFNQWTLNYSGALTNNGLAPPVAPT 420  
DB 361 PKLGSVVLGTWEDRDFINQPARFPTVGLYTDHFNQWALPNYSGALTNNGLAPPVAPT 420  
QY 421 FPGEQILFRSHIPLKGGVADPVIDCLLPQEWIQHLYQESAPSDQDVALIRFTNPDTCRV 480  
DB 421 FPGEQILFRSHIPLKGGVADPVIDCLLPQEWIQHLYQESAPSDQDVALIRFTNPDTCRV 480  
QY 481 LFEAKLHRSQGIITVANTGSRPIVVPANGYFRDFTWVWQFYSLAPMGNGRRVQ 535  
DB 481 LFEAKLHRSQGIITVANTGSRPIVVPANGYFRDFTWVWQFYSLAPMGNGRRVQ 535  
RESULT 14  
Q91V51 PRELIMINARY; PRT; 535 AA.  
ID Q91V51  
AC Q91V51  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus NLV/Ludwigslust/218/99/DE.  
OC Viruses; serNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OC NCBI\_taxid=122910;  
RX [1]  
RC SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Girlington/93/UK;  
RX MEDLINE=20404883; PubMed=10949950;  
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,  
RA Clegg J.C., Chamberlain J., Brown D.W.G.;  
RT "Capsid protein diversity among 'Norwalk-like' viruses.";  
RL Virus Genes 20:227-236(2000).  
DR EMBL; AJ277606; CAB89087.1;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 535 AA; 58772 MW; 27C6C7DBFC2A6AB CRC64;  
Query Match 86.2%; Score 2452; DB 12; Length 535;  
Best Local Similarity 84.3%; Pred. No. 3.7e-189;  
Matches 451; Conservative 37; Mismatches 47; Indels 0; Gaps 0;  
QY 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60  
DB 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60  
QY 61 NGEFTVSPRNSPQEVILNLELGPENLPLAHLRSMYNGYAGGVQVLLAGNAFTAGKLV 120  
DB 61 NGEFTVSPRNSPQEVILNLELGPENLPLAHLRSMYNGYAGGVQVLLAGNAFTAGKLV 120  
QY 121 FAAPDPHPLENI SFGQITMFPVHIIDVRLTEPVLPLPDVRNFFHYNQNEPRMLVA 180  
DB 121 FAAPDPHPLENI SFGQITMFPVHIIDVRLTEPVLPLPDVRNFFHYNQNEPRMLVA 180  
QY 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNFLVLPETVESKTPFTLPILTIGELTNS 240  
DB 181 MLYTPLRNSGDDVFTVSCRVLTRPSDFDNFLVLPETVESKTPFTLPILTIGELTNS 240



Db 181 MLYTPLRNSGDDVFTVSCRVPSPDFDFNYLVPPTVESKTPFTLPILTIGELNS 240  
Qy 241 RPPVPIDELYSNPSLVQPQNGCALDGELOGTTLQTLPTAICSPRGRINQKVSGENHV 300  
Db 241 RPPVPIDELYSNPSLVQPQNGRTLLDGELOGTTLQTLVPSNICSRGRINAHLSNQHR 300  
Qy 301 WNMQVNTINGTPTDPTGDPVAPLGTDFSGKLVGVLSQRDHNAACRSHDAVIATNSAKPT 360  
Db 301 WNMQVNTANGTPTDPTEDVPALGTDFGLANTYGVTSQRNPONTCAHGDGILATWSKPT 360  
Qy 361 PKLGAIQGTWEEDVHINQKFTPVGLFENEGNQWTLPNYSGALTINMGLAPPVAPT 420  
Db 361 PKLGSVVLGTWEEDRDINQPTFRFPVGLYDTHNQWALPNYSGALTINMGLAPVAPL 420  
Qy 421 FPEGQILFFRSHIPKGGVADVIDCLLPQEWIQHLYQESAPSDVALIRFTNPTDGRV 480  
Db 421 FPEGQILFFRSHIPKGGTSNGAIDCLLPQEWVQHYQESAPSDVALIRFTNPTDGRV 480  
Qy 481 LFEAKLHRSYITVANTGSRPIVWPANGYFRDFTWVNOFYSLAPMGTGNGRRRVQ 535  
Db 481 LFEAKLHRCGFTVANTGSRPIVWPANGYFRDFTWVNOFYSLAPMGTGNGRRRVQ 535

RESULT 15

Q915C7 PRELIMINARY; PRT; 548 AA.  
AC Q915C7;  
DT 01-DEC-2001 (TREMELrel. 19, Created)  
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)  
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus NLV/Exfurt/546/00/DE.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=173929;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NLV/Exfurt/546/00/DE;  
RA Kuenkel U., Hoehne M., Schreier E.;  
RT "Molecular epidemiology of outbreaks of gastroenteritis associated  
RT with Norwalk-like viruses in Germany.";  
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF427118; AAL18874.1;  
DR InterPro; IPR004005; Calici.coat.  
DR Pfam; PF00915; Calici.coat; 1.  
SQ SEQUENCE 548 AA; 59874 MW; 9A1D8BEE58ADF7D CRC64;

Query Match 81.2%; Score 2307.5; DB 12; Length 548;  
Best Local Similarity 77.7%; Pred. No. 1.7e-177;  
Matches 426; Conservative 57; Mismatches 52; Indels 13; Gaps 3;  
Qy 1 MKMASNDAAPSNDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60  
Db 1 MKMASNDAAPSNDGAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60  
Qy 61 NGEFTVSPRNSPCEVLLNLELPELNPYLALSRMYNGYAGGVQVLLAGNAFTAGKLV 120  
Db 61 NGEFTVSPRNSPCEVLLNLELPELNPYLALSRMYNGYAGGVQVLLAGNAFTAGKLV 120  
Qy 121 FAAPVPHPLENISQGITWTFPHVVIDVRLTLPVLLPDVRRNFFHYNQONEPRMLVA 180  
Db 121 FAAPVPHPLENISQGITWTFPHVVIDVRLTLPVLLPDVRRNFFHYNQONEPRMLVA 180  
Qy 181 MLYTPLRNSGDDVFTVSCRVLTRPSPDFDFNYLVPPTLESKTPFTLPILTIGELTNS 240  
Db 181 MLYTPLRNSGDDVFTVSCRVLTRPSPDFDFNYLVPPTLESKTPFTLPILTIGELTNS 240  
Qy 241 RPPVPIDELYSNPSLVQPQNGCALDGELOGTTLQTLPTAICSPRGRINQKVSGENH- 299  
Db 241 RPPVLPIDVLYTNPESAIVQCNCRCTLDGELQGTTLQTLPTGICAPRGKVTQVQDEHRG 300

Qy 300 -VWMQVNTINGTPTDPTGDPVAPLGTDFSGKLVGVLSQRDHN-----ACRSHDA 350  
Db 301 THMNTVTLNAGTTPDPTEDVPALGTDFSGIYGVISQRNTNTVPGEGLPANRAHEA 360  
Qy 351 VIATNSAKFTPKLGAIQGTWEEDVHINQKFTPVGLFE---NEGFNQWTLPNYSGAL 407  
Db 361 VIATYSPKFTPKLGNIOFSTWETQDVSSGQPTKFTPVGLASVDANSHDFDQWTLPSYSGAL 420  
Qy 408 TLNMGLAPPVAPTFFGEQILFFRSHIPKGGVADVIDCLLPQEWIQHLYQESAPSDV 467  
Db 421 TLNMGLAPPVAPTFFGEQILFFRSHIPKGGVADVIDCLLPQEWIQHLYQESAPSDV 480  
Qy 468 ALIRFTNPTDGRVLFEAKLHRSYITVANTGSRPIVWPANGYFRDFTWVNOFYSLAPMGT 527  
Db 481 ALVRYNPETGRTLFEAKLHRCGFTVANTGSRPIVWPANGYFRDFTWVNOFYSLAPMGT 540  
Qy 528 GNGRRRVQ 535  
Db 541 GSGRRRMQ 548

Search completed: June 1, 2004, 13:53:20  
Job time : 32.1485 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time, 46.5611 Seconds  
(without alignments)  
3289.030 Million cell updates/sec

Title: US-09-926-799-9  
Perfect score: 2863  
Sequence: 1 MKMASNDAAPSNDGAASLVP.....VNQFYSLAPMTGTGNGRRRIQ 542

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2863	100.0	542	4	AAB49708	Small rou
2	2048.5	71.6	535	4	AAB49707	Small rou
3	1933	67.5	540	4	AAB49706	Small rou
4	1931	67.4	548	4	AAB49705	Small rou
5	1886	65.9	548	5	AAU91272	Norwalk v
6	1872	65.4	550	4	AAB49709	Small rou
7	1821.5	63.6	539	4	AAB49704	Small rou
8	1538.5	53.7	541	4	AAB49710	Small rou
9	1181	41.3	530	4	AAB49701	Small rou
10	1179	41.2	530	2	AAAS0972	Norwalk v
11	1179	41.2	530	7	ADC72176	Norwalk v
12	1172.5	41.0	545	4	AAB49700	Small rou
13	1165.5	40.7	544	4	AAB49703	Small rou
14	1155.5	40.4	546	4	AAB49702	Small rou
15	1137	39.7	530	2	AAAS7091	Small rou
16	336	11.7	579	2	AAW08143	RHDV caps
17	279	9.7	668	2	AAAR10686	Feline ca
18	279	9.7	668	4	AAE04304	Feline ca
19	278	9.7	623	4	AAB47044	Feline Ca
20	278	9.7	623	4	AAB47043	Feline Ca
21	276.5	9.7	547	4	AAW50108	Feline ca
22	276.5	9.7	671	4	AAW50107	Feline ca
23	275	9.6	622	4	AAB47045	Feline Ca
24	265.5	9.3	668	4	AAB67462	Amino aci
25	265.5	9.3	669	4	AAB67461	Amino aci

26	181	6.3	40	5	AAB49708	Standard; protein; 542 AA.
27	168	5.9	40	5	AAB49707	
28	149.5	5.2	934	1	AAB20016	
29	145.5	5.1	1147	5	ABB76724	
30	133	4.6	6310	6	ABU39869	
31	119	4.2	2206	2	AAR22210	
32	116	4.1	394	7	ABE65406	
33	114.5	4.0	594	6	ABU37475	
34	112.5	3.9	3930	6	ABU18893	
35	111	3.9	572	4	ABG07638	
36	111	3.9	639	4	ABG94220	
37	110.5	3.9	1165	7	ADC00875	
38	110.5	3.9	6304	6	ABU09236	
39	109.5	3.8	568	4	ABB64772	
40	109.5	3.8	993	1	AAP50231	
41	109.5	3.8	2227	1	AAP60066	
42	109.5	3.8	2468	6	ABU38411	
43	109.5	3.8	2468	6	ABP59933	
44	108.5	3.8	854	1	AAP50287	
45	108.5	3.8	1077	2	AAW95559	

ALIGNMENTS

RESULT 1  
AAB49708  
ID AAB49708 standard; protein; 542 AA.  
AC AAB49708;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 9.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 95JP-00175928.  
XX  
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PI (DENK-) DENKA SEIKEN KK.  
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
WPI: 2001-080848/09.  
N-PSDB; AAF29149.  
Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.  
Claim 1; Page 59-61; 84pp; Japanese.  
This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

Query Match 100.0%; Score 2863; DB 4; Length 542;  
Best Local Similarity 100.0%; Pred. No. 2.2e-237;  
Sequence 542 AA;

Matches	542;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MMKASNDAA	PSNDGAASLVPEGINETMPLPVPVAGASIAAPVAGQTNIIIDPWIRTNFVQAP	60					
Db	1	MMKASNDAA	PSNDGAASLVPEGINETMPLPVPVAGASIAAPVAGQTNIIIDPWIRTNFVQAP	60					
QY	61	NGEFTVSP	RNSPGEILLNLELGPDLNPNYLAHLRSMYNGYAGGVEVQVLLAGNAFTAGKIL	120					
Db	61	NGEFTVSP	RNSPGEILLNLELGPDLNPNYLAHLRSMYNGYAGGVEVQVLLAGNAFTAGKIL	120					
QY	121	FAAIPN	FLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFHFNNQPPQMRLLVA	180					
Db	121	FAAIPN	FLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFHFNNQPPQMRLLVA	180					
QY	181	MLYTP	LRNSGDDVFTVSCRVLTRPTDFEYILVPPSVESKTKPFTLPILITISLTNS	240					
Db	181	MLYTP	LRNSGDDVFTVSCRVLTRPTDFEYILVPPSVESKTKPFTLPILITISLTNS	240					
QY	241	RFPIE	IQLYTAPNETNVVQCNGRCCTLDGELQGTTLSSAVCFIQGRTVADNGDNDQ	300					
Db	241	RFPIE	IQLYTAPNETNVVQCNGRCCTLDGELQGTTLSSAVCFIQGRTVADNGDNDQ	300					
QY	301	NLQLTY	PNPGASVPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYISITS	360					
Db	301	NLQLTY	PNPGASVPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYISITS	360					
QY	361	GKTPK	IGSLHSITTEHVHPNQSRFTPVGVAVDENTFPQWVLPHYAGSLALNTNLAP	420					
Db	361	GKTPK	IGSLHSITTEHVHPNQSRFTPVGVAVDENTFPQWVLPHYAGSLALNTNLAP	420					
QY	421	AVAPT	PEQEQLLFFRSRVPVQGLQGDADFIDCLLPQEWVNHFYQEAAPSQADVALIRV	480					
Db	421	AVAPT	PEQEQLLFFRSRVPVQGLQGDADFIDCLLPQEWVNHFYQEAAPSQADVALIRV	480					
QY	481	NPDT	GRTLFEAKLHRSFITYVSHGTGAYPLVPPNGHFRFDSWVNFYSLAPMGTGNRRR	540					
Db	481	NPDT	GRTLFEAKLHRSFITYVSHGTGAYPLVPPNGHFRFDSWVNFYSLAPMGTGNRRR	540					
QY	541	IQ	542						
Db	541	IQ	542						
RESULT 2									
ID	AA49707								
XX	AA49707	standard; protein; 535 AA.							
XX	AA49707;								
XX	04-APR-2001	(first entry)							
XX	Small round structured virus protein SEQ ID 8.								
XX	Small round structured virus; SRSV; food poisoning.								
XX	Small round structured virus.								
XX	WO200079280-A1.								
XX	28-DEC-2000.								
XX	22-JUN-2000; 2000WO-JP004095.								
XX	22-JUN-1999; 99JP-00175928.								
XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.								
XX	(DENK-) DENKA SEIKEN KK.								
XX	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;								
XX	WPI; 2001-080848/09.								
XX	N-PSDB; AAF29148.								

PT	Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.
PT	Claim 1; Page 57-59; 84pp; Japanese.
XX	This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks
XX	Sequence 535 AA;
QY	Query Match 71.6%; Score 2048.5; DB 4; Length 535;
Db	Best Local Similarity 71.7%; Pred. No. 2.9e-167; Indels 17; Gaps 7;
QY	Matches 392; Conservative 64; Mismatches 74;
Db	1 MKVASNDAA
QY	61 NGEFTVSP
Db	61 NGEFTVSP
QY	121 FAALP
Db	121 FAALP
QY	181 MLYTP
Db	181 MLYTP
QY	241 RFPIE
Db	241 RFPIE
QY	298 WQNL
Db	301 MN
QY	357 STTGK
Db	353 ATNSAK
QY	416 TNLAP
Db	411 MGLAP
QY	476 LIRYV
Db	469 LIRF
QY	536 NGRRIQ
Db	529 NGRRRVQ
RESULT 3	
ID	AA49706
XX	AA49706 standard; protein; 540 AA.
XX	AA49706;
XX	04-APR-2001 (first entry)
XX	Small round structured virus protein SEQ ID 7.
XX	Small round structured virus; SRSV; food poisoning.

OS Small round structured virus.  
XX WO200079280-A1.  
XX PD 28-DEC-2000.  
XX PF 22-JUN-2000; 2000WO-JP004095.  
XX PR 22-JUN-1999; 99JP-00175928.  
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX PA (DENK-) DENKA SEIKEN KK.  
XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX DR N-PSDB; AAF29147.  
XX XX  
PT Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX  
XX Claim 1; Page 54-57; 84pp; Japanese.  
XX  
CC This invention relates to a kit for the detection and typing of small  
CC round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX  
XX Sequence 540 AA;  
Query Match 67.5%; Score 1933; DB 4; Length 540;  
Best Local Similarity 67.0%; Pred. No. 2.6e-157; Mismatches 94; Indels 10; Gaps 6;  
Matches 366; Conservative 76; Mismatches 94; Indels 10; Gaps 6;  
QY 1 MKWASNDAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNTIIPDPIRTNFVQAP 60  
DB 1 MKWASNDATPSNDGAAGLVPESSNEAMALEPVVGLASLAAPVGTQNTIIPDPIRTNFVQAP 60  
QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVNGYAGGVVQVLLAGNAFTAGKIL 120  
DB 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVNGYAGGVVQVLLAGNAFTAGKII 120  
QY 121 FAATPPNPLVDMSPAQITMLPHLIVDVRTLEPIMTLPDVRNVFVHNNQPPQMRLLVA 180  
DB 121 FAATPPNPLVDMSPAQITMLPHLIVDVRTLEPIMTLPDVRNVFVHNNQPPQMRLLVA 180  
QY 181 MLYTPLRNSGDDVFTVSCRVLTPTDPDFEYILVPPSVESKTKPFTLPILTITSELNLS 240  
DB 181 MLYTPLRNSGDDVFTVSCRVLTPTDPDFEYILVPPSVESKTKPFTLPILTITSELNLS 240  
QY 241 RFPPIEQLYTAPNETNVVQCGNCRCTLDGELQGTQLLSSAVCFLOGRT---VADNGN 297  
DB 241 RFPPIEQLYTAPNETNVVQCGNCRCTLDGELQGTQLLSSAVCFLOGRT---VADNGN 297  
QY 298 WQONLLQLYTPNGASYDPTDEVPAPLCTQDFSGMLYGVLTQDNVNVSTGEAKNAGIYIS 357  
DB 301 WQONLLQLYTPNGASYDPTDEVPAPLCTQDFSGMLYGVLTQDNVNVSTGEAKNAGIYIS 357  
QY 358 TTSKFTPKIGSIGHLSITEHVHPNQOSRFTPVGV-AVDENTPFOQWVLPHYAGSLALNT 416  
DB 357 TTSKFTPKIGSIGHLSITEHVHPNQOSRFTPVGV-AVDENTPFOQWVLPHYAGSLALNT 416  
QY 417 NLAPAVAPTPFGEQLLFFRSVPCVQGLQGDADFICLLPQSWNVHFQEAAPSQADVAL 476  
DB 417 NLAPAVAPTPFGEQLLFFRSVPCVQGLQGDADFICLLPQSWNVHFQEAAPSQADVAL 476  
QY 477 IRYVNPDTGRTLFEAKLHRSIGITVSHTCAYPLVPPNGHFRFDSWVNFQFYSLAPMGTCN 536  
DB 475 IRYVNPDTGRTLFEAKLHRSIGITVSHTCAYPLVPPNGHFRFDSWVNFQFYSLAPMGTCN 536

QY 537 GRRRIQ 542  
DB 535 GRRRVQ 540

RESULT 4  
AAB49705  
ID AAB49705 standard; protein; 548 AA.

XX AC AAB49705;  
XX DT 04-APR-2001 (first entry)  
XX DE Small round structured virus protein SEQ ID 6.  
XX KW Small round structured virus; SRSV; food poisoning.  
XX OS Small round structured virus.  
XX PN WO200079280-A1.  
XX PD 28-DEC-2000.  
XX PF 22-JUN-2000; 2000WO-JP004095.  
XX PR 22-JUN-1999; 99JP-00175928.  
XX PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
XX PA (DENK-) DENKA SEIKEN KK.  
XX PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX DR WPI; 2001-080848/09.  
XX DR N-PSDB; AAF29146.  
XX PT Kit for the detection and typing of small round-structured virus (SRSV)  
XX PT strains for investigation of food poisoning outbreaks, contains  
XX PT antibodies.  
XX PS Claim 1; Page 52-54; 84pp; Japanese.  
XX CC This invention relates to a kit for the detection and typing of small  
XX CC round structured virus (SRSV) strains. The kit contains antibodies  
XX CC directed against peptides represented in sequences AAB49700 - AAB49710,  
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
XX CC used for detecting and typing strains of SRSV in order to prevent the  
XX CC spread of infection and to examine the epidemiology of outbreaks  
XX  
XX Sequence 548 AA;

Query Match 67.4%; Score 1931; DB 4; Length 548;  
Best Local Similarity 65.7%; Pred. No. 3.9e-157; Mismatches 102; Indels 18; Gaps 5;  
Matches 364; Conservative 70; Mismatches 102; Indels 18; Gaps 5;

QY 1 MKWASNDAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNTIIPDPIRTNFVQAP 60  
DB 1 MKWASNDAPSNDGAAGLVPESSNEAMALEPVVGLASLAAPVGTQNTIIPDPIRTNFVQAP 60  
QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVNGYAGGVVQVLLAGNAFTAGKIL 120  
DB 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVNGYAGGVVQVLLAGNAFTAGKII 120  
QY 121 FAATPPNPLVDMSPAQITMLPHLIVDVRTLEPIMTLPDVRNVFVHNNQPPQMRLLVA 180  
DB 121 FAATPPNPLVDMSPAQITMLPHLIVDVRTLEPIMTLPDVRNVFVHNNQPPQMRLLVA 180  
QY 181 MLYTPLRNSGDDVFTVSCRVLTPTDPDFEYILVPPSVESKTKPFTLPILTITSELNLS 240  
DB 181 MLYTPLRNSGDDVFTVSCRVLTPTDPDFEYILVPPSVESKTKPFTLPILTITSELNLS 240  
QY 241 RFPPIEQLYTAPNETNVVQCGNCRCTLDGELQGTQLLSSAVCFLOGRT---RTVADN 294

```

|||||:|:|:| | ||||| ||||| ||||| :|:|:| |:|
241 RFPVPIESLHTSFTENIVQCGRVTLLDGLMGTTLQLPSQICAFRGVLTRSTSRASDQ 300

295 GDNWDONL-----IQLTYPNGASVDPTDEVPAPLGTDQDFSGMLYGVLTQDNVNVSITGEA 348
      ||| ||| ||| :|:|:| ||| ||| ||| :|:|:| ||| ||| :|:|:| ||| |||
301 ADTATPRLFNYYWHVQDLNMGTPYDPAEDIPGFLGTFDRGKVFVGASQRNLID-STTRA 359

349 KNAKGIIYSTSKGFTPKIGSIGLSHISITEHVHPNQOGRFPTPGVAVDENTPFQQWVLPHY 408
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
360 HEAK---VDTTAGRTFKPLGSLEISTDSDDFDQNQPKTFPVGLGVGDNEAEFCQWSLPDY 416

409 AGSLALNTNLAPAVAPTFFEGEQLLFRRSRVPCVOGLOGQDAFIDCLLPQEWNHFYOEA 468
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
417 SGQTHNMNLAPAVAPNFPEQEQLLFRRSQLPSSGG--RSNGVLDCLPQEWVHQFYOSA 474

469 PSQADVALIRYNPDGRTLFEAKLHRSGFTIVSHTGAYPLVPPNGHGFDSWVNQFYS 528
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
475 PAQTQVALRVYNPDFTKVLFEAKLHGKFMTIANNGDSPITVPDPNGYFRFESWVNPFFT 534

529 LAPMGTGNRRRIQ 542
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
535 LAPMGTGNRRRIQ 548

RESULT 5
AAU91272
D AAU91272 standard; protein; 548 AA.
X AAU91272;
T 18-JUN-2002 (first entry)
X Norwalk virus associated polynucleotide #1.
E Norwalk virus; monoclonal antibody; geno group I; geno group II;
X Nowalk virus; monoclonal antibody; geno group I; geno group II;
X immunological detection; food; viral infection.
X Norwalk virus.
X JP2002020399-A.
X 23-JAN-2002.
X 10-JUL-2000; 2000JP-00208151.
X 10-JUL-2000; 2000JP-00208151.
XX (OSAP ) OSAKA PREFECTURE.
XX (NISE-) NIPPON SEIBUTSU KAGAKU CENT KK.
XX (IATR ) IATRON LAB INC.
XX WPI; 2002-287412/33.
XX A monoclonal antibody useful in the immunological detection and diagnosis
XX of Norwalk virus infection.
XX Disclosure; Page 12-13; 24pp; Japanese.
XX The invention describes a monoclonal antibody recognising Norwalk virus,
XX a capsid protein of Norwalk virus, or a common antigen epitope on the
XX capsid protein molecule of geno group I and geno group II. The antibody
XX is useful for immunological detection and quantitative analysis of
XX Norwalk virus in foods and the serum of infected patients. This sequence
XX represents a Norwalk virus associated protein described in the invention
XX Sequence 548 AA;
SQ

Query Match 65.9%; Score 1886; DB 5; Length 548;
Best Local Similarity 65.3%; Pred.No.2.9e-153;
Matches 359; Conservative 66; Mismatches 107; Indels 18; Gaps

Oy 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNIIDPWIRTNFVOAP 60

```

PS Claim 1; Page 62-64; 84pp; Japanese.

XX This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX Sequence 550 AA;

Query Match 65.4%; Score 1872; DB 4; Length 550;  
 Best Local Similarity 64.9%; Pred. No. 4.7e-152;  
 Matches 361; Conservative 74; Mismatches 101; Indels 20; Gaps 7;

QY 1 MKMASNDAAPSNDGAASIVPEGINETMPLEPVAGASIAAPVAGQNIIDPWIRTFVQAP 60  
 DB 1 MKMASNDAAPSNDGAASIVPEANLEPVAGASIAAPVAGQNIIDPWIRTFVQAP 60

QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGGVEVQVLLAGNAFTAGKIL 120  
 DB 61 QGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGGVEVQVLLAGNAFTAGKII 120

QY 121 FFAAIPNPLVDMISPAQITMLPHLIVDVRTLEPIMTPLPDVNRVYHFNNQPPMRLLVA 180  
 DB 121 FFAAIPNPLVDMISPAQITMLPHLIVDVRTLEPIMTPLPDVNRVYHFNNQPPMRLLVA 180

QY 181 MLYTPLRSNGSGDDVFTVSCRVLTPTDPEFIYLVPPSVESKTPFTLITISELTNS 240  
 DB 181 MLYTPLRSNGSGDDVFTVSCRVLTPTDPEFIYLVPPSVESKTPFTLITISELTNS 240

QY 241 RPIPIEOLYAPNETNVPVQCGRCCTLDGELQGTTLSSAVCFLOQRTVADN---GDN 297  
 DB 241 RPIPIEOLYAPNETNVPVQCGRCCTLDGELQGTTLSSAVCFLOQRTVADN---GDN 297

QY 298 WQNLLOLTYP-----NGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGE-- 347  
 DB 298 WQNLLOLTYP-----NGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGE-- 347

QY 348 -AKNAKGIYSTCKFTKIGSLHSITEHVENQOSRFTPVGVAVDENTPFOOWLP 406  
 DB 348 -AKNAKGIYSTCKFTKIGSLHSITEHVENQOSRFTPVGVAVDENTPFOOWLP 406

QY 407 HYAGSLALNTLAPAVTFFGEOQLLFFRSRVPVCGQGGQDAFIDCLLPQEWVNHFYQE 466  
 DB 407 HYAGSLALNTLAPAVTFFGEOQLLFFRSRVPVCGQGGQDAFIDCLLPQEWVNHFYQE 466

QY 467 AAPSQADVALIRYVNPDTGRTLFEAKLHRSFGITVSHGTGAYPLVPPNGHFRFDSWVNOF 526  
 DB 467 AAPSQADVALIRYVNPDTGRTLFEAKLHRSFGITVSHGTGAYPLVPPNGHFRFDSWVNOF 526

QY 527 YSLAPMGTCNGRRRIQ 542  
 DB 527 YSLAPMGTCNGRRRIQ 542

QY 535 YTLAPMGSCQGRRAQ 550  
 DB 535 YTLAPMGSCQGRRAQ 550

RESULT 7  
 AAB49704  
 ID AAB49704 standard; protein; 539 AA.  
 XX  
 XX AAB49704;  
 AC  
 DT 04-APR-2001 (first entry)  
 XX  
 DE Small round structured virus protein SEQ ID 5.  
 XX  
 KW Small round structured virus; SRSV; food poisoning.  
 XX  
 OS Small round structured virus.  
 XX  
 PN WO200079280-A1.

PD 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.

XX 22-JUN-1999; 95JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

PA (DENK-) DENKA SEIKEN KK.

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

DR MPI; 2001-080848/09.

DR N-PSDB; AAF29145.

XX Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.

XX Claim 1; Page 50-52; 84pp; Japanese.

CC This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks  
 XX Sequence 539 AA;

Query Match 63.6%; Score 1821.5; DB 4; Length 539;  
 Best Local Similarity 63.7%; Pred. No. 1e-147;  
 Matches 347; Conservative 74; Mismatches 111; Indels 13; Gaps 6;

QY 1 MKMASNDAAPSNDGAASIVPEGINETMPLEPVAGASIAAPVAGQNIIDPWIRTFVQAP 60  
 DB 1 MKMASNDAAPSNDGAASIVPEANLEPVAGASIAAPVAGQNIIDPWIRTFVQAP 60

QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGGVEVQVLLAGNAFTAGKIL 120  
 DB 61 QGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGGVEVQVLLAGNAFTAGKII 120

QY 121 FFAAIPNPLVDMISPAQITMLPHLIVDVRTLEPIMTPLPDVNRVYHFNNQPPMRLLVA 180  
 DB 121 FFAAIPNPLVDMISPAQITMLPHLIVDVRTLEPIMTPLPDVNRVYHFNNQPPMRLLVA 180

QY 181 MLYTPLRSNGSGDDVFTVSCRVLTPTDPEFIYLVPPSVESKTPFTLITISELTNS 240  
 DB 181 MLYTPLRSNGSGDDVFTVSCRVLTPTDPEFIYLVPPSVESKTPFTLITISELTNS 240

QY 241 RPIPIEOLYAPNETNVPVQCGRCCTLDGELQGTTLSSAVCFLOQRTVADNNDQ 300  
 DB 241 RPIPIEOLYAPNETNVPVQCGRCCTLDGELQGTTLSSAVCFLOQRTVADNNDQ 300

QY 301 NLLQTYTPNGASDYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYSTTS 360  
 DB 301 NLLQTYTPNGASDYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYSTTS 360

QY 361 CKFTPKIGSLHSITEH-VHPNQOSRFTPVGVAVD-----ENTPQOWVLPVHAGSLALN 415  
 DB 361 CKFTPKIGSLHSITEH-VHPNQOSRFTPVGVAVD-----ENTPQOWVLPVHAGSLALN 415

QY 416 TNLAPAVAPTFFGEOQLLFFRSRVPVCGQGGQDAFIDCLLPQEWVNHFYQAAAPQADVA 475  
 DB 416 TNLAPAVAPTFFGEOQLLFFRSRVPVCGQGGQDAFIDCLLPQEWVNHFYQAAAPQADVA 475

QY 476 LIRYVNPDTGRTLFEAKLHRSFGITVSHGTGAYPLVPPNGHFRFDSWVNOF 535  
 DB 476 LIRYVNPDTGRTLFEAKLHRSFGITVSHGTGAYPLVPPNGHFRFDSWVNOF 535

QY 536 NGRRR 540  
 DB 536 NGRRR 540

QY 533 AGRRR 537  
 DB 533 AGRRR 537





QY 56 FVQAPNGEFTVSPRNSPGEILLNLBELGDLNPLYLALHLSRMVNGYAGGVEVQVLLAGNAPT 115  
 Db 60 FVQAPNGEFTVSPRNSPGEILLNLBELGDLNPLYLALHLSRMVNGYAGGVEVQVLLAGNAPT 119  
 QY 116 AGKILFAAPNPNFLVDMISPAQITMLPHLIVDVRTLEPMTPLDVRNVFVHFNNQOPQR 175  
 Db 120 AGKIIVSCIPPGFGSHNLTIQAATLFPHVIAVRTLDPIEVLDPVNRVLFHNNDRNOOT 179  
 QY 176 MRLVAMLYTPLRNSGDDVFTVSCRVLTRPTDPDEFIYLVPPSVESKTKPFTLPILTIS 235  
 Db 180 MRLVAMLYTPLRNSGDDVFTVSCRVLTRPTDPDEFIYLVPPSVESKTKPFTLPILTIS 239  
 QY 236 ELTNSRFPPIPIEQLYTAPNETNVQCGRCCTLDGELQTTQLLSSAVCFLOGRTVADNG 295  
 Db 240 SLNSRPLPISSMGISPDNVQVQFNGRCCTLDGRLVGTTPVSLSHVAKIRG---TSNG 296  
 QY 296 DNWDQNLQLTYPNCGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNNVNSTGEAKNAGIY 355  
 Db 297 -----TVINLTDELDTGTPHFF-EGPAPIGFPLGDCW-----HINMTQFGHSSQOTQYD 344  
 QY 356 ISTTSKFTPKIGSI---GLHS-----ITEHVHPNQQRFTVPVAVADENTPFGQWVLP 407  
 Db 345 VDTTPTDFVPHLGSIQANGISGNVGVLSWSPESH-----PSGQVD-----LWKIPN 394  
 QY 408 YAGSLALNTNLAPAVAPTFPGEQLLFFRSRVPVCGVQLOQDAP-IDCLLPQEWVNHFYOE 466  
 Db 395 YGSSITEATHLAPSVVPPGFGVLFVFFMSKIP-----GPGAYSLPCLLPQVISHLASE 448  
 QY 467 AAPSQADVALIRYVNPDTGRTLFELAKLHRSGETITVSHTGAY--PLVVPNGHFRPDSWVN 524  
 Db 449 QAPTVEAALLHYVDPDTGRTLGEKAYPDGELTCVPNGASSGPOQLPINGVFVFSWVS 508  
 QY 525 QFSLAPMGNGRR 539  
 Db 509 RPYQLKPVGTASSAR 523

## RESULT 10

AARS0972

ID AARS0972 standard; protein; 530 AA.

XX AARS0972;

AC AARS0972;

XX 16-OCT-2003 (revised)

DT 25-MAR-2003 (revised)

DT 05-OCT-1994 (first entry)

XX Norwalk virus strain 8FIIa protein (encoded by ORF2).

DE Norwalk virus; pathogen; acute gastroenteritis; food poisoning;

XX Norwalk virus; pathogen; acute gastroenteritis; food poisoning;

KW seafood contamination; diagnostic assay; calicivirus; small round virus.

XX Norwalk virus; (strain 8FIIa).

OS Norwalk virus; (strain 8FIIa).

XX WO9405700-A2.

XX 17-MAR-1994.

XX 07-SEP-1993; 93WO-US008447.

XX 07-SEP-1992; 92US-00941365.

XX (BAYU ) BAYLOR COLLEGE MEDICINE.

XX Matson DO, Estes MK, Jiang X, Graham DY;

XX WPI; 1994-101125/12.

XX N-PSDB; AAQ56826.

XX DNA from Norwalk and related viruses - used for preparing prods. for use

PT in diagnostic assays, detection and vaccines for Norwalk and related

PT viruses.

XX

PS Claim 14; Page 68-70; 156pp; English.

XX

CC The Norwalk virus was isolated from stool samples from adult volunteers  
 CC infected with safety tested Norwalk virus strain 8FIIa. The coding  
 CC sequence is useful for the design of probes for use in diagnostic assays  
 CC for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct  
 CC PN field.) (Updated on 16-OCT-2003 to standardise OS field)

XX

SQ Sequence 530 AA;

Query Match 41.2%; Score 1179; DB 2; Length 530;

Best Local Similarity 46.5%; Pred. No. 2.1e-92;

Matches 256; Conservative 84; Mismatches 171; Indels 40; Gaps 12;

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

QY

Db

```

PD 03-JUN-2003.
XX
PF 07-JUN-1995; 95US-00486049.
XX
XX 08-NOV-1989; 89US-00433492.
PR 27-APR-1990; 90US-00515993.
PR 27-AUG-1990; 90US-00573509.
PR 06-MAY-1991; 91US-00696454.
XX
XX (BAYU ) BAYLOR COLLEGE MEDICINE.
PA
XX
XX Estes MK, Jiang X, Graham DY;
PI
XX
XX WPI; 2003-776005/73.
DR N-PSDB; ADC72174.
DR
XX
XX Inducing an immune response against non-Norwalk virus agents, comprises
PT administering an immunogen recombinantly expressed from a cDNA from
PT Norwalk virus.
XX
XX
XX Example 4; SEQ ID NO 3; 45pp; English.
XX
XX This invention relates to a novel method of inducing an immune response
CC in an individual against Norwalk virus and non-Norwalk virus agents, by
CC orally or parenterally administering an immunogen recombinantly expressed
CC or synthesized from a cDNA of Norwalk virus given in the specification.
CC Norwalk virus is one of the most important viral pathogens, causing acute
CC gastroenteritis. The invention may be used for the development of
CC compounds with virucidal activity or an antiviral vaccine. The present
CC sequence is the amino acid sequence of a protein encoded by the Norwalk
CC virus genome of the invention.
XX
XX
XX Sequence 530 AA;
XX
XX
XX Query Match 41.2%; Score 1179; DB 7; Length 530;
XX Best Local Similarity 46.5%; Pred. No. 2.1e-92;
XX Matches 256; Conservative 84; Mismatches 171; Indels 40; Gaps 12;
XX
XX 1 MKMASNDAPSDGAA---SLVPGINETWPL--EPVAGASIAAPVAGOTNIIDPWRTN 55
XX 1 MMMAASKDATSSVDGASGAGQLVPE-VNASDPLAMDVPAGSTAVATAGQNPIDPWINN 59
XX
XX 56 FVOAPNGEFTVSPRNSPGEILLNLELGPDLNPLPYLAHLSRMVYAGGVQVLLAGNAFT 115
XX 60 FVQAPQGEFTISPNTPGDVFLDLSLGPLNPLFLLHLSQMYNGVGNVRIMLAGNAFT 119
XX
XX 116 AGKILFAAIPNPLVDMTSPAQITMLPHLIVDVRTLEPIMTPLPDVNRVYFHNQOPQR 175
XX 120 AGKIIVSCIPPGFGSHNLTAQAATLFPFHVIADVRLDPIEVPLEDVRNVLHNNDRNQOT 179
XX
XX 176 MRLVAMLYTPLRNSGDDVFTVSCRVLTPTPPEFIYLVPPSVESKTKPFTLPILTIS 235
XX 180 MRLVCMLYTPLTGGTGDSFVAGRWMTCPSPDENFLFVPTVEQTRPFTLPLNPLS 239
XX
XX 236 ELTNSRFPIPIPOLYATARNVTVQNGRCRTLDGELQGTQTLLSSAVCFLOQRTVADNG 295
XX 240 SILNSRAPLPISSMGISPDNVQVQFQNGRCRTLDGRLVGTTPVSLSHVAKIRG---TSNG 296
XX
XX 296 DNWDQNLLQLTYPNGASYDPTDEVPAPLQTQDFSCMLYXGVLTDQNNVNVSTGEAKNAGIY 355
XX 297 -----TVINLTDLGTDPFHPF-EGPAPIGPDLGGCDW-----HINTQGHSSQTYD 344
XX
XX 356 ISTTSCKFTPKGISGLHSITEHVHVNQOSRFT----PVGVAVDENTPFQOVLPHYAGS 411
XX 345 VDTTPTTFVPHLGSIQANGISGNYVGVLSWISPSHPSGSQVD-----LWKIPNYGSS 398
XX
XX 412 LALNTLAPAVAPTFPGEOLLFPRSRVPCVQLOGODAF-IDCLLPQEWNVHFYGEAAPS 470
XX 399 ITEATHLAPSVYPGGEVLVFFMFSKMP-----GPGAYNLCPQLPQETISHLASEQAPT 452
XX
XX 471 QADVALIRYVNPDTGRTLPEAKLHRGGFITVSHGTAY--PLVVPNGHFRPDSWVNYQFYS 528
XX 453 VGEAALLHYVDPTGRTNLGEFKAYPDGFLTCVNGASSGPPQLPFGVVFVSVSRFYQ 512

```

```

QY 529 LAPMGTGNGRR 539
Db 513 LKPVGTASSAR 523

RESULT 12
AAB49700
ID AAB49700 standard; protein; 545 AA.
XX
XX AAB49700;
XX AC
XX 04-APR-2001 (first entry)
XX
XX Small round structured virus protein SEQ ID 1.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX WO200079280-A1.
XX
XX 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
XX N-PSDB; AAF29141.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX
XX Claim 1; Page 40-42; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710, -
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 545 AA;
XX
XX Query Match 41.0%; Score 1172.5; DB 4; Length 545;
XX Best Local Similarity 46.4%; Pred. No. 8e-92;
XX Matches 257; Conservative 85; Mismatches 185; Indels 27; Gaps 15;
XX
XX 1 MKMASNDAPSN---DGAASLVPE-GINETMPLFPVAGASIAAPVAGOTNIIDPWRTN 55
XX 1 MMMAASKD-APTNWDGTSGAGQLVPEANTAEPISEMFVAGATAATAATAGVNMIDPWIMN 59
XX
XX 56 FVOAPNGEFTVSPRNSPGEILLNLELGPDLNPLPYLAHLSRMVYAGGVQVLLAGNAFT 115
XX 60 FVQAPQGEFTISPNTPGDVFLDLSLGPLNPLFLLHLSQMYNGVGNVRIMLAGNAFT 119
XX
XX 116 AGKILFAAIPNPLVDMTSPAQITMLPHLIVDVRTLEPIMTPLPDVNRVYFHNQOPQR 175
XX 120 AGKIIVSCIPPGFGSHNLTAQAATLFPFHVIADVRLDPIEVPLEDVRNVLHNNDRNQOT 178
XX
XX 176 MRLVAMLYTPLRNSG--SGDDVFTVSCRVLTPTPPEFIYLVPPSVESKTKPFTLPILT 233
XX 179 MRLVCMLYTPLRASGSSGTDPFVIAGRVLTCPSPDFSLFLVPPNVQKTFKFSVPNLP 238
XX 234 ISELTNSRFPIPIEOLYTAPEVQNGRCRTLDGELQGTQTLLSSAVCFLOQRTVAD 293

```

```

Db 239 LMTLSNSRVSILKSMWVRDGMQVQFQNGRVTLDDGQGGTTPTASOLCKIRGSVFHA 298
QY 294 NGDNWDQNLQLTYPNNGASYDPTDEVPAPLGTQDFSGMLYGLVLTQDNVNVSTGEAKNAG 353
Db 299 NGGN-GYNLTJEL---DGSFYH-AFESPAIGFPDLGECMDHMEASPTTQFNTGDV--IKQ 351
QY 354 IYVSTSGKFTPKIGSI---GLHSITEHVHPNQQRFTPVGVAVD-ENTPFOQWVLPHYA 409
Db 352 INVQESA-FAPHLGTIQADGLSDVS--VNTNMIAKLGWVSPVSDCHRGDVPWIPRYG 408
QY 410 GSLALNTNLAPAVATFPCEQLLFRSRVPCVQGLQGQDAFIDCLLPQSWNHFFQEARP 469
Db 409 SILTEAAQLAPPIYPPGFEALVFFMSDPPIAHGNTGLS--VPCTIPQEFVTHFVNEQAP 466
QY 470 SQADVALIRYVNPDTGRTLFEAKLHRSGETIV--SHTGAYPLVVPNGHFRDQSWNOFY 527
Db 467 TRGEAALLHYLPDTHRNLFGEFLYPEGFWTCVPNSSGTGPTLPINGVVFVSWSRFY 526
QY 528 SLAPMGTGNRRRI 541
Db 527 QLKPVGTAGPACRL 540

```

## RESULT 13

```

ID AAB49703
XX AAB49703 standard; protein; 544 AA.
AC AAB49703;
XX
XX 04-APR-2001 (first entry)
DT Small round structured virus protein SEQ ID 4.
DE Small round structured virus; SRSV; food poisoning.
KW Small round structured virus.
XX Small round structured virus.
OS
XX WO200079280-A1.
PN
XX 28-DEC-2000.
PD
XX 22-JUN-2000; 2000WO-JP004095.
PF
XX 22-JUN-1999; 99JP-00175928.
PR
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI: 2001-080848/09.
DR N-PSDB; AAF29144.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
XX Claim 1; Page 47-49; 84pp; Japanese.
PS
XX
XX This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 544 AA;
SQ

```

Query Match 40.7%; Score 1165.5; DB 4; Length 544;

Best Local Similarity 45.3%; Pred. No. 3.2e-91;

Matches 256; Conservative 89; Mismatches 171; Indels 49; Gaps 16;

```

QY 1 MKAASNDAAPSND---GAASLVPEGIN--ETWPLEPVAGASIAAPVAGQTNIIDPWIRN 55
Db 1 MMASKDATPSADGATGAGQLVPE-VNTADPIPIDVAGSSTALATAGQNLIDPMIINN 59
QY 56 FVQAPNGEFTVSPRNSPGSIIILNLELGDPLNPNYLAHLSRMVNGYAGGVEQVLLAGNAFT 115
Db 60 FVQAPQGEFTISPNNTPGDVLPDLQLGPHLNPFLSHLSQMYNGWGNMRVRVVLAGNAFT 119
QY 116 AGKILFAAIPNPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVVYHFNQOPQR 175
Db 120 AGKVIICCVPCFQFQRTLSIAQATLFPHVIAADVRLDPEVPLEDVRNVLYH-NNDTQPT 178
QY 176 MRLVAMLYTPLRNSGS--GDDVFTVSCRVLTRPTPDFEFIYLVPPSVSKTKPFTLPILT 233
Db 179 NELLCLMYLPLRTGASGCTDSFVAVRVLTCPGPDFNFLFLVPPTEQKTRPFTVPNIP 238
QY 234 ISELNTPRPPIPIEOLYTPAPNETNVVQNGRECTLDGELQGTQLLSSAVCFLOGRTVAD 293
Db 239 LKYLNSRIPNPIEGMSLSPDQTNVQFQNGRECTIDGQPLGTPTPVSVSOLCKFRGRITS- 297
QY 294 NGDNWDQNLQLTYPNNGASYDPTDEVPAPLGTQDFSGMLYGV-----LTQDNVNVST 345
Db 298 -----GQVNLNLTLDGSPF-WAPAAPAPAGFPDLGSCDWHIEMSKI PNSTQNPVTN 351
QY 346 GEAKNAKGIYITTSKFTPKIGSIGL----HSITEHVHPNQQRFTPVGVAVDENTPFO 401
Db 352 S-----VKPNSQQFVPHLSITLDBENVSSGGDYIGTIQWT--SPPSDSGGANTNF- 399
QY 402 QWVLPHYAGSLALNTNLAPAVATFPGEQLLFRSRVPCVQGLQGQDAFIDCLLPQSWNV 461
Db 400 -WKIPDYGSSLAESAQLAPAVYPGFEVIVYFMAVIP-GPNQSGSPNLVPCLLPQEYIT 457
QY 462 HFYQEAAPSQADVALIRYVNPDTGRTLFEAKLHRSGETIV-----SHTGAYPLVVPNGHF 517
Db 458 HFISEQAIQGEAALLHYVDPDTNRNLGEFLYPCGYLTCVPNSSSTG--PQQLPLDGVF 515
QY 518 RFDQSWNOFYSLAPMGT-GNGRRRI 541
Db 516 VFASWVSRYQLKPVGTAGPARGRL 540

```

## RESULT 14

```

ID AAB49702
XX AAB49702 standard; protein; 546 AA.
AC AAB49702;
XX
XX 04-APR-2001 (first entry)
DT Small round structured virus protein SEQ ID 3.
DE Small round structured virus; SRSV; food poisoning.
KW Small round structured virus.
XX Small round structured virus.
OS Small round structured virus.
XX WO200079280-A1.
PN
XX 28-DEC-2000.
PD
XX 22-JUN-2000; 2000WO-JP004095.
PF
XX 22-JUN-1999; 99JP-00175928.
PR
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI: 2001-080848/09.
DR N-PSDB; AAF29143.
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT

```

diagnostic assay; human calicivirus; small round virus; SRSV; KY89;  
Norwalk virus; capsid protein.

Small round structured virus.

WO9405700-A2.

17-MAR-1994.

07-SEP-1993; 93WO-US008447.

07-SEP-1992; 92US-00941365.

(BAYU ) BAYLOR COLLEGE MEDICINE.

Matson DO, Estes MK, Jiang X, Graham DY;

WPI, 1994-101125/12.

N-PSDB; AAQ56832.

DNA from Norwalk and related viruses - used for preparing prods. for use in diagnostic assays, detection and vaccines for Norwalk and related viruses.

Example 7; Fig 13a; 156pp; English.

The known sequence for Norwalk virus was used to obtain the sequence of other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide cDNA sequence includes part of the polymerase region and the capsid region of the genome; the deduced amino acid sequences are AAR57092 and AAR57091, respectively. Expression of fragments and derivs. of Norwalk-related viruses permits development of diagnostic assays to detect antibodies, antigens, viral genetic material or antivirals. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS field.)

Sequence 530 AA;

Query Match 39.7%; Score 1137; DB 2; Length 530;

Best Local Similarity 45.8%; Pred. No. 8.7e-89;

Matches 254; Conservative 85; Mismatches 168; Indels 48; Gaps 16;

1 MKMASNDAAPSNDGAA--SLVPEGIN--ETMPLEPVAGASIAAPVAGQTNIIDPWIRTN 55

1 MMASKDATSVDDGASASQVLPE-VNASDPLAMDVPAGSSTAVATAGQVNPIDPWINN 59

56 FVOAPNGEFTVSPRNSPGEILLNLELGPDLNPLVLAHLSRMVNGYAGGVEVLLAGNAFT 115

60 FVOAPQGEFTISPNTPGDVLFDLSLGPDLNPLVLAHLSRMVNGYAGGVEVLLAGNAFT 119

116 AKGILFAAIPNPLVDMSIPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQOPQR 175

120 AGKIIIVCSIPPGFGSQQLTIAQTLFPHVIADVRTLDFIEVPLEDVRNVLFHNDNQOT 179

176 MRLVAMLYTLPLRSNGSGDDVFTVSCRVLTRPTDPFEFYLVPSPVESKTKPFTLPIITIS 235

180 MRLVCMLYTLPLSTGGTGDSFVVGAVMTCPSPDENFLFVPTVEQTRPFTLENPLS 239

236 ELTNSRFPPIPIEQLYTAPNETNVQCONGRCTLDGELQGTQQLSSAVCFLQGRVADNG 295

240 SLNSRAPLPIGMIISPDNVQSVQFQNGRCTLGRLVTTPTVSLSHVAKING---TNG 296

296 DNDWQNLQLQTYPNNGASYPDTPVAPLQTDQFSGMLYGVLTQDNNVNSTGBAKNAKGIY 355

297 ----TVINLTDLGTPPHFF-EGPAPIGFPLDGGCDW-----HINMTQFGHSSQOTYD 344

356 ISTTSKFTPKIGSI---GLHS-----ITEHVHNPQOSRFTPVGVAVDENTFFQQVLP 407

345 VDTTPDTSVPHLGSIQANGIGSGNYIGVLSWSPSPSH---PSGSDV-----LWKIN 394

408 YAGSLAINTLAPAV-APTFFGQLLFFRSRVPVQGLQCADFIDCLLPQEWVNFHYOE 466

PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.

Claim 1; Page 45-47; 84pp; Japanese.

This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

XX Sequence 546 AA;

Query Match 40.4%; Score 1155.5; DB 4; Length 546;

Best Local Similarity 45.3%; Pred. No. 2.3e-90; Indels 53; Gaps 18;

Matches 258; Conservative 88; Mismatches 171;

1 MKMASNDAAPSNDGAA--SLVPEGIN--ETMPLEPVAGASIAAPVAGQTNIIDPWIRTN 55

1 MMASKDATSVDDGASAGQVLPE-VNTADPLPMEVPAGFTTAVATAGQVNMIDPWVNN 59

56 FVOAPNGEFTVSPRNSPGEILLNLELGPDLNPLVLAHLSRMVNGYAGGVEVLLAGNAFT 115

60 FVQSPQGEFTISPNTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMRVIRLLAGNAPS 119

116 AKGILFAAIPNPLVDMSIPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQOPQR 175

120 AGKIIIVCSIPPGFGTSSSLTIAQTLFPHVIADVRTLEPIEMPLEDVRNVLYHTNDN-QPT 178

176 MRLVAMLYTLPLRS-NGSGD--DVFTVSCRVLTRPTDPFEFYLVPSPVESKTKPFTLPIIT 233

179 MRLVCMLYTLPLRTGGSGNSDSFVVGAVMTCPSPDENFLFVPTVEQTRPFTLENPL 238

234 ISELTNSRFPPIPIEQLYTAPNETNVQCONGRCTLDGELQGTQQLSSAVCFLQGRVAD 293

239 LQTLNSRFPPIPIEQLYTAPNETNVQCONGRCLIDQGLTTPATSGQLFRVYRGI--- 295

294 NGDNWQNLQLQTYPNNGASYPDTPVAPLQTDQFSGMLYGV-LTQDNNVNSTGBAKNAK 352

296 ---NQARTLNLTEVDGKPFMAFDS-PAPVGFDPFGKCDWHMRISKTPNNTSSGDP--MR 349

353 GIYISTSGKFTPKIGSIGLHSITEHVHP-----NQOSRFTPVGVAVDENTPFG 401

350 SVSVQTNVQGFVPHLGSIQFDEVFN--HPTGDYGTIEWISQPS---TPPG-----TDIN 399

402 QWVLPHYAGSLALNTNLAPAVAPTTPGQQLFFRSRVPVQGLQGQDA--FIDCLLPQEW 459

400 LWEIPDYGSSLSQAANLAPVFPFGFGEALYFVSAPF---GPNRSAPNDVPCLLPQEF 456

460 VNHFYQEAAPSQADVALTRYNPDTGRTLFEAKLHRSQFIV--SHTGAYPLVVPNGHF 517

457 ITHFVSEQAPTMGDAALLHYDPTNRLGEBFKLYPGGYLTCVPGVAGPQQLPLNGVF 516

518 RDSWNVQFSLAPMGTGN-----GRRRI 541

517 LFVSWVSRYQLKPGVTASTARSRLQVRRI 546

RESULT 15

AAR57091 standard; protein; 530 AA.

XX AAR57091;

XX 27-AUG-2003 (revised)

XX 25-MAR-2003 (revised)

XX 05-OCT-1994 (first entry)

XX Small round virus SRSV/KY/89 capsid protein.

XX pathogen; acute gastroenteritis; food poisoning; seafood contamination;

Db 395 YGSSITEATHLAPSVSPGF-GEVLVFFMSKIPG-----PGGDS-LPCLLPQGYISHLASE 448  
Qy 467 AAPSQADVALIRYVNPDTGRTLFEAKLHRSGETVTSHTGAY--PLVVPENGHFRFDSWYN 524  
Db 449 QAPTVEGEGFLLHYVDPDIDRNLGEFKAYPDGFLTCVPNGCASSGPQQLPINGVVFVSWYS 508  
Qy 525 QFYSLAPMGNGRR 539  
Db 509 RFYQLKPVGTASTAR 523

Search completed: June 1, 2004, 13:45:59  
Job time : 47.5611 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.9134 Seconds  
(without alignments)  
2166.837 Million cell updates/sec

Title: US-09-926-799-9  
Perfect score: 2863  
Sequence: 1 MKMASNDAPNSDGAASLVP.....VNQFYSLAPMTGNGRRRIQ 542

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA\*  
1: /cgn2\_6/ptodata/2/iaa/5A-COMB.pep:\*  
2: /cgn2\_6/ptodata/2/iaa/5B-COMB.pep:\*  
3: /cgn2\_6/ptodata/2/iaa/6A-COMB.pep:\*  
4: /cgn2\_6/ptodata/2/iaa/6B-COMB.pep:\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS-COMB.pep:\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1179	41.2	530	US-08-486-049-3	Sequence 3, Appli
2	278	9.7	623	US-09-590-020-2	Sequence 2, Appli
3	278	9.7	623	US-09-590-020-4	Sequence 4, Appli
4	275	9.6	622	US-09-590-020-6	Sequence 6, Appli
5	272.5	9.5	626	US-09-590-020-7	Sequence 7, Appli
6	265.5	9.3	668	US-09-617-594A-4	Sequence 4, Appli
7	265.5	9.3	669	US-09-617-594A-2	Sequence 2, Appli
8	115	4.0	2206	US-07-852-260-2	Sequence 2, Appli
9	115	4.0	2206	US-08-461-503-2	Sequence 2, Appli
10	115	4.0	2206	US-08-465-250-2	Sequence 2, Appli
11	108.5	3.8	1091	5516630-2	Patent No. 5516630
12	108.5	3.8	2227	US-08-475-886-2	Sequence 2, Appli
13	108.5	3.8	2227	US-08-475-886-4	Sequence 4, Appli
14	108.5	3.8	2227	US-08-475-886-6	Sequence 6, Appli
15	108.5	3.8	2227	US-08-397-232-2	Sequence 2, Appli
16	108.5	3.8	2227	US-08-397-232-4	Sequence 4, Appli
17	108.5	3.8	2227	US-09-171-387-2	Sequence 2, Appli
18	108.5	3.8	2227	US-09-653-499-2	Sequence 2, Appli
19	108.5	3.8	2227	US-09-653-499-4	Sequence 4, Appli
20	108.5	3.8	2227	US-09-653-499-6	Sequence 6, Appli
21	108.5	3.8	2227	US-10-104-966-12	Sequence 12, Appli
22	108.5	3.8	2227	US-10-135-988-2	Sequence 2, Appli
23	108.5	3.8	2227	US-10-135-988-4	Sequence 4, Appli
24	108.5	3.8	2227	US-10-135-988-6	Sequence 6, Appli
25	108	3.8	2318	US-09-091-219-24	Sequence 24, Appli
26	108	3.8	2318	US-09-660-541-24	Sequence 24, Appli
27	107.5	3.8	2736	US-09-252-991A-30227	Sequence 30227, A

28	106.5	3.7	740	4	US-10-164-595-6	Sequence 6, Appli
29	106.5	3.7	755	4	US-10-164-595-10	Sequence 10, Appli
30	106.5	3.7	781	4	US-10-164-595-4	Sequence 4, Appli
31	104.5	3.7	762	4	US-09-107-532A-4275	Sequence 4275, Ap
32	103.5	3.6	839	1	US-08-087-016-2	Sequence 2, Appli
33	101.5	3.5	1257	1	US-08-340-428B-49	Sequence 49, Appli
34	100.5	3.5	496	4	US-10-164-595-8	Sequence 8, Appli
35	100.5	3.5	496	4	US-10-164-595-60	Sequence 60, Appli
36	99.5	3.5	1056	4	US-09-079-030-217	Sequence 217, App
37	99.5	3.5	4536	4	US-09-180-422B-27	Sequence 27, Appl
38	99.5	3.5	4536	4	US-09-079-030-1	Sequence 1, Appli
39	99.5	3.5	4563	4	US-09-108-006C-1	Sequence 1, Appli
40	98.5	3.4	1638	4	US-09-071-035-258	Sequence 258, App
41	98.5	3.4	1638	4	US-09-071-035-262	Sequence 262, App
42	98.5	3.4	1638	4	US-09-071-035-266	Sequence 266, App
43	96.5	3.4	579	6	5223424-13	Patent No. 5223424
44	95.5	3.3	834	4	US-09-252-991A-31321	Sequence 31321, A
45	95	3.3	426	4	US-09-107-532A-7172	Sequence 7172, Ap

ALIGNMENTS

RESULT 1  
US-08-486-049-3  
; Sequence 3, Application US/08486049  
; Patent No. 6572862  
; GENERAL INFORMATION:  
; APPLICANT: Bates, Mary K  
; APPLICANT: Jiang, Xi  
; APPLICANT: Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE:  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-486-049-3

Query Match 41.2%; Score 1179; DB 4; Length 530;  
Best Local Similarity 46.5%; Pred. NO. 5e-107;  
Matches 256; Conservative 84; Mismatches 171; Indels 40; Gaps 12;  
QY 1 MKMASNDAPNSDGAASLVP-----SLVPEGINETMPL--EPVAGASIAAPVAGQTHLIDPWITN 55  
DB 1 MMASKDSTSDVGASGAGQLVPE-VNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIIN 59

127 NPLVDMISPAQITML-----PHLIVDVRTLEPIMTLPDVRNVFVHNNQPOQRMRLVAMLY 183  
194 G-----IEFVQSTMLQYPHVLFDFARQVEPVIFAIPDRSNLXHLMSD-TTSLVIMVY 247  
184 TPLRNSGSDVFTVSC--RVLTPTPDFFIYLVPP-----SVESK----- 223  
248 NDL--INPYANDTNSGGCIYVETKPGDFKPHLLKPPGSMITHGVSFSDLPKSSSLWIG 306  
224 -----TKPFTL-----PI-LTISELTNSRFPPIPIBQLY 250  
307 NRHWSIDITFIIRPFVFOANRHFDFNQETAGWSTPRPRPITITVSESNMKLGIGVATDY 366  
251 TAPN-----ETNV-----VQCNGR--CTLDG-----ELQTTOLLSSAVCF 286  
367 IYVGPIDGPDPTIIEQLTTPAGIYISITASNGSVITTAAGYDAAETIVNTTNFKSMYIC-- 424  
287 QGRTVADNDNDQNLQIT--YPMGASYDPTDEV-----PAPLG---TODPSGM 331  
425 -GSLQRAWGDKKISNTAFITTAVRKNSIEPSNTIDMTKLVVYQDAHVGEVQTSIDILA 483  
332 LYGVLTQDNNVNSTGEAKNAG-----IYISTTSKFTPKIGSIGLHSTIEHVHPNOQS 385  
484 LLGY-----TGIEEAIIGSDRDKVVRISV---LPETGARGGN-----HP1FYK 523  
386 RFTPVGVAVDENTPFOQWVLPHYAGSLANTNLAP 420  
524 NSIKLGVIRSIDVFNSQL-HTSRQLSLNNYLLP 557

RESULT 3  
US-09-590-020-4  
; Sequence 4, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; PRIOR FILING DATE: 2000-06-08  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 4  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
; US-09-590-020-4

Query Match 9.7%; Score 278; DB 4; Length 623;  
Best Local Similarity 25.0%; Pred. No. 2.6e-18;  
Matches 129; Conservative 65; Mismatches 177; Indels 144; Gaps 26;

12 NDGAASLVPEGINETMP-----LEPVAGASIAAPVAGQTNIIIDPW-IRTNFVQAPNGEFTV 66  
81 DDGSITTPEQG---TWVGVIAPPSAQMSAAADMATKGSVDSEWAEFFSFTSVNW---- 133  
67 SPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKILFAAIP 126  
134 STSETQKILFKQSLGPLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKILFAAIP 193  
127 NPLVDMISPAQITML-----PHLIVDVRTLEPIMTLPDVRNVFVHNNQPOQRMRLVAMLY 183  
194 G-----IEFVQSTMLQYPHVLFDFARQVEPVIFAIPDRSNLXHLMSD-TTSLVIMVY 247  
184 TPLRNSGSDVFTVSC--RVLTPTPDFFIYLVPP-----SVESK----- 223  
248 NDL--INPYANDTNSGGCIYVETKPGDFKPHLLKPPGSMITHGVSFSDLPKSSSLWIG 306  
224 -----TKPFTL-----PI-LTISELTNSRFPPIPIBQLY 250

56 FVQAPNGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFT 115  
60 FVQAPNGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFT 119  
116 AKKILFAAIPNPLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFVHNNQPOQR 175  
120 AGKIIIVSCIPPGFGHNLTIAQATLPHVIAVDRVLTLEPIMTLPDVRNVFVHNNQPOQR 179  
176 MELVAMLYTPLRNSGSDVFTVSCRVLTPTPDFFIYLVPPSVESKTKFTLPIITIS 235  
180 MELVAMLYTPLRNSGSDVFTVSCRVLTPTPDFFIYLVPPSVESKTKFTLPIITIS 239  
236 ELTNSRFPPIPIBQLYTAPNETNVVQNGRCTLDELOQTOLLSSAVCFQGRVADNG 295  
240 SLNSRAPLPISSMGISPDNVQVQNGRCTLDELOQTOLLSSAVCFQGRVADNG 296  
296 DNDWQNLQITYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNNVNSTGEAKNAGIY 355  
297 -----TVINLTDELGTGPHFF-EGPAPIGFDPDLGGCDW-----HINMTQFGHSSQTYD 344  
356 ISTTSKFTPKIGSIGLHSTIEHVHPNOQSRT-----PVGAVDENTPFOQWVLPHYAGS 411  
345 VDTPTDFVPHLGSIOANGISGNGVGLSVISPPSHPSGSQVD-----LWKIPNYGSS 398  
412 LALNTNLAPAVATFPGEQQLFFRSRVPVCGQLQGDADF-IDCLLPQEWNVHFOEAPAS 470  
399 ITEATHLAPSVYPPGFEVLVFFMSKWP-----GFGAYNLPCLLPQEVISHLASEQAPT 452  
471 QADVALIRYVNPDTGRTLFEAKLHRSQFTVSHGTAY--PLVVPNGHFRFDSWNOFYS 528  
453 VGEAALLHYVDPDTGRNLGEFKAYPDGFLTCVPGNASSGQQLPINGVFVFSWVSRRFYQ 512  
529 LAPMGTGNRR 539  
513 LKPVGTASSAR 523

RESULT 2  
US-09-590-020-2  
; Sequence 2, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; PRIOR FILING DATE: 2000-06-08  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: Patent In Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 623  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
; US-09-590-020-2

Query Match 9.7%; Score 278; DB 4; Length 623;  
Best Local Similarity 25.0%; Pred. No. 2.6e-18;  
Matches 129; Conservative 65; Mismatches 177; Indels 144; Gaps 26;

12 NDGAASLVPEGINETMP-----LEPVAGASIAAPVAGQTNIIIDPW-IRTNFVQAPNGEFTV 66  
81 DDGSITTPEQG---TWVGVIAPPSAQMSAAADMATKGSVDSEWAEFFSFTSVNW---- 133  
67 SPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKILFAAIP 126  
134 STSETQKILFKQSLGPLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKILFAAIP 193



Db 307 NRHWSIDTPIIRPFVQANRHFDFNQETAGSTRFRFRITITVSNNKSLGIGVATDY 366  
QY 251 TAPN-----ETNV-----VQCONGR--CTLDG-----ELQGTTLSSAVCFPL 286  
Db 367 IVPGIPDGWPDITIPQLTPAGIYSITASNGTDITTAAGYDAAEIVTNTNFKSMYIC-- 424  
QY 287 QGRTVADNGDNDQNLLOIT--YPNGASYDPTDEV-----PAPLG-----TQDFSGM 331  
Db 425 -GSLQRAWGDKKISNTAFITTAVRKGNSEIPEPSNTIDMTKLAVVYQDAHVGEVQTSIDTILA 483  
QY 332 LYGVLTDQNNVNVSTGEAKNAGK-----IYISTSGKFTPKIGSIGLHSHITEHVHPNQOS 385  
Db 484 LLOY-----TGIGBEAIGSDRDKVVRISV-----LPETGARGGN-----HPIFYK 523  
QY 386 RFTPVGVADVENTPFQOWVLPFYAGSLAINTNLAP 420  
Db 524 NSIKLGVIRSIDVFNSQL-HTSRQLSLNNYLLP 557  
RESULT 4  
US-09-590-020-6  
; Sequence 6, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; PRIOR FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 6  
; LENGTH: 622  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-6

Query Match 9.6%; Score 275; DB 4; Length 622;  
Best Local Similarity 24.6%; Pred. No. 5.1e-18;  
Matches 126; Conservative 62; Mismatches 185; Indels 140; Gaps 24;  
QY 12 NDGAASLVPEGINETMP-----LEPVAGASIAAPVAGQTNIIIDPW-IRTNFVQAPNGEFTV 66  
Db 81 DGSITTPBQG---TWVGGVIAEPSQAQMAAADMATGKSVDSWEAEFFSFTSVNM---- 133  
QY 67 SPRNSPGEILLNLELGPDLNPLYLAHLRMVNGVAGVEVQVLLAGNAFTAGKILFAAIPP 126  
Db 134 STSETQKILFKQSLGPLNPLYSLAKUYVALAGSEVRFSGVFGGKLAIVVP 193  
QY 127 NFLVDMISPAQITML---PHLIVDVRTLEPIMPLPDVRNVFYHFNQPPQRMRLVAMLY 183  
Db 194 G-----IEPVQSTMLQYPHVLFDAQVPEVIFAIPDLASNLVHMSDITD-TTSLVIMVY 247  
QY 184 TPLRNSGDDVFTVSC--RVLTPTPDPDFEYILVPP-----SVESK----- 223  
Db 248 NDL-INPYANDNTSSGCIIVTVETKPGDPFKFLLKPPGSMITHGSPVSDLPKSSSLMTIG 306  
QY 224 -----TKPFTL-----PI-LTISELTNSRFPPIEQLY 250  
Db 307 NRHWSIDTPIIRPFVQANRHFDFNQETAGSTRFRFRITITVSNNKSLGIGVATDY 366  
QY 251 TAPN-----ETNV-----VQCONGR--CTLDG-----ELQGTTLSSAVCFPL 286  
Db 367 IVPGIPDGWPDITIPQLTPAGIYSITASNGTDITTAAGYDAAEIVTNTNFKSMYICGS 426  
QY 287 QGRTVADNGDNDQNLLOIT-----YPNGASYDPTDEV-----PAPLG-----TQD 327

Db 427 LQRA-----WDDKKISNTAFITTAVRKGNSEIPEPSNTIDMTKLAVVYQDAHVGEVQTSID 479  
QY 328 FSGMLYGVLTQDNNVNVSTGEAKNAGKIYISTSGKFTPKIGSIGLHSHITEHVHPNQOSRF 387  
Db 480 ITLALLGY-----TGIGBEAIG-----SDRDKVVRISVLGETGARGGNHPIFYKNS 525  
QY 388 RFTPVGVADVENTPFQOWVLPFYAGSLAINTNLAP 420  
Db 526 IKLGVIRSIDVFNSQL-HTSRQLSLNNYLLP 557  
RESULT 5  
US-09-590-020-7  
; Sequence 7, Application US/09590020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnis, Alvars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; PRIOR FILING DATE: 2000-06-08  
; PRIOR APPLICATION NUMBER: 60/138,484  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-7

Query Match 9.5%; Score 272.5; DB 4; Length 626;  
Best Local Similarity 23.2%; Pred. No. 9.2e-18;  
Matches 137; Conservative 84; Mismatches 211; Indels 159; Gaps 30;  
QY 11 SNDGAASLVPEGINETMP-----LEPVAGASIAAPVAGQTNIIIDPW-IRTNFVQAPNGEFT 65  
Db 80 ADDGSITAPBQG---TWVGGVIAEPSQAQMAAADMATGKSVDSWEAEFFSFTSVNM--- 133  
QY 66 VSPRNSPGEILLNLELGPDLNPLYLAHLRMVNGVAGVEVQVLLAGNAFTAGKILFAAIPP 125  
Db 134 -STSETQKILFKQSLGPLNPLYSLAKUYVALAGSEVRFSGVFGGKLAIVVP 192  
QY 126 NFLVDMISPAQITMLPHLIVDVRTLEPIMPLPDVRNVFYHFNQPPQRMRLVAMLYTP 185  
Db 193 PG--VDPVOSTMLQYPHVLFDAQVPEVIFCLPDLRSTLYHMSDITD-TTSLVIMVYND 249  
QY 186 LRSNGSGDDVFTVSC--RVLTPTPDPDFEYILVPP-----SVESK----- 223  
Db 250 L-INPYANDNTSSGCIIVTVETKPGDPFKFLLKPPGSMITHGSPVSDLPKSSSLWGNR 308  
QY 224 -----TKPFTL-----PI-LTISELTNSRFPPIEQLYTA 252  
Db 309 HWSIDTPIIRPFVQANRHFDFNQETAGSTRFRFRITITVSNNKSLGIGVATDY 368  
QY 253 PN-----ETNVQCONGRCTLDGELQGTTLSSAVCFLOGRTVADNGDNDQNLLOIT 306  
Db 369 PGIPDGWPDIT-----TIPGEL-----IPAGDYAITNGTGND-----IT 401  
QY 307 YPNGASYDPTDEVPAPLTQDFSGM-LYGVLTQDNNVNVSTGEAKNAGKIYIST-----T 359  
Db 402 TATG--YTDALIK---NNTNFRGMVYICSLQR-----AWGDKKISNTAFITATLDGDN 451  
QY 360 SKGFTPKIGSIGLHSHITEHVHPNQOSR-----FTPVG---VAVDENTPFQOWV 404  
Db 452 NNKINP-CNTIDQSKIVVQDAHVGVKKAQTSDDTLALLGYTGIGRQAIGSDRDRVRIST 510  
QY 405 LPHYAGSLAINTNLAPAVAPTFPGEQQLLFFPRSRVPCVOGLOQDAFIDCLL---PQEWVN 461  
Db 511 LPETGAR-----GGNHPIFYKNSIKLGVIRSIDVFNSQLHSHITEHVHPNQOSLN 555



```

; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/852,260
; FILING DATE: 19920619
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-B-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 977-9550
; TELEFAX: (212) 664-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: AMINO ACID
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-07-852-260-2

Query Match 4.0%; Score 115; DB 1; Length 2206;
Best Local Similarity 18.7%; Pred. No. 0.23;
Matches 105; Conservative 65; Mismatches 179; Indels 212; Gaps 25;

QY 3 MASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVA-----GQTNIIDP- 50
Db 284 LAIDSMVKHNNWGIALP-----LSPDFAQDSSVEIPITVTIAPMCSEFNGLRNVTPAK 338
QY 51 -----WIRTNFVQAPNG--EFTVSPR-NSPGEI----- 75
Db 339 FQGLPVLNTPGSNQVLTSDNHQSPCAIPEFDVTPIDIPGEVKNMELAEIDTMIPLNLE 398
QY 76 -----LNLLELGPDLNPYLAH-----LSRMVNGVAGGVQV 107
Db 399 STKRTMDMYRVTLSDSADLSQPICLSLSPAFDPRLSHTMTLGEVLNYTHWAGSLKFTF 458
QY 108 LLAGNAFTAGKILFAAIPNPLVDIMISPAQITMLPHLIVDVRTLPEINTPLPDVNRVYH 167
Db 459 LFCGSMMATGKILVAYAPPG-AQPPTSKEAMLGTHVLDLGLQSSCTWVVPWISNVYR 517
QY 168 FNNQPO-----PRMLVAMLYTPLRNSGDDVFTVSCRVLTRPTDPFEIYL-- 215
Db 518 OTTQDSFTEGGYISMFYQTRIVVPLSTPKMSMLG---FVSACN-----DFSRLLRD 567
QY 216 -----VPPVESKTKPFTLPILTISELTNSRFPPI-----IEQL 249
Db 568 TTHISQSALPQGIEDLTSEVAQGALTLS-LPKQDLSLPTDKASGPAHSKEVPALTAVETG 626
QY 250 YT---APNET---NVQCON-----GRCT---LDGE----- 271
Db 627 ATNPLAPSDTVQTRHVQRRSRSESTIESFFARGACVAIIIEVDNEQPTTRAQKLFAWMRI 686
QY 272 -LQGTITQLLSSAVCFLOGR-----TVADNGDNWDQNLQLTY-PNGASYDPT 316
Db 687 TYKDTVQLRRKLEFFTYSRFDMFTFVVYNTANNGHALNQ-VYQIMYIPPGA----- 740
QY 317 DEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYSTTSKFTPKIGSIGLHSIT 376
Db 741 ---PTPKSWDDY-----TWQTSNPSIFTYTGAAPARISVYPVGLANAY 781
QY 377 EHVHPNQOSRRFTPVGVAVDEN 397
Db 782 SHFYDG----FAKVPLKTDAN 798

RESULT 9
US-08-461-503-2
; Sequence 2, Application US/08461503
; Patent No. 5834302
; GENERAL INFORMATION:
; APPLICANT: Racaniello, Vincent
; APPLICANT: Tatem, Joanne M.
```

```

; APPLICANT: Weeks-Levy, Carolyn L.
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES
; TITLE OF INVENTION: FROM CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10112
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,503
; FILING DATE: 5-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-D-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX: 422523 COOP UI
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-503-2

Query Match 4.0%; Score 115; DB 2; Length 2206;
Best Local Similarity 18.7%; Pred. No. 0.23;
Matches 105; Conservative 65; Mismatches 179; Indels 212; Gaps 25;

QY 3 MASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVA-----GQTNIIDP- 50
Db 284 LAIDSMVKHNNWGIALP-----LSPDFAQDSSVEIPITVTIAPMCSEFNGLRNVTPAK 338
QY 51 -----WIRTNFVQAPNG--EFTVSPR-NSPGEI----- 75
Db 339 FQGLPVLNTPGSNQVLTSDNHQSPCAIPEFDVTPIDIPGEVKNMELAEIDTMIPLNLE 398
QY 76 -----LNLLELGPDLNPYLAH-----LSRMVNGVAGGVQV 107
Db 399 STKRTMDMYRVTLSDSADLSQPICLSLSPAFDPRLSHTMTLGEVLNYTHWAGSLKFTF 458
QY 108 LLAGNAFTAGKILFAAIPNPLVDIMISPAQITMLPHLIVDVRTLPEINTPLPDVNRVYH 167
Db 459 LFCGSMMATGKILVAYAPPG-AQPPTSKEAMLGTHVLDLGLQSSCTWVVPWISNVYR 517
QY 168 FNNQPO-----PRMLVAMLYTPLRNSGDDVFTVSCRVLTRPTDPFEIYL-- 215
Db 518 OTTQDSFTEGGYISMFYQTRIVVPLSTPKMSMLG---FVSACN-----DFSRLLRD 567
QY 216 -----VPPVESKTKPFTLPILTISELTNSRFPPI-----IEQL 249
Db 568 TTHISQSALPQGIEDLTSEVAQGALTLS-LPKQDLSLPTDKASGPAHSKEVPALTAVETG 626
QY 250 YT---APNET---NVQCON-----GRCT---LDGE----- 271
Db 627 ATNPLAPSDTVQTRHVQRRSRSESTIESFFARGACVAIIIEVDNEQPTTRAQKLFAWMRI 686
QY 272 -LQGTITQLLSSAVCFLOGR-----TVADNGDNWDQNLQLTY-PNGASYDPT 316
Db 687 TYKDTVQLRRKLEFFTYSRFDMFTFVVYNTANNGHALNQ-VYQIMYIPPGA----- 740
QY 317 DEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYSTTSKFTPKIGSIGLHSIT 376
```

Db 518 QTTQDSFTEGGYISMFYQTRIVVPLSTPKSNMLG---FVSACN-----DFSRLLRD 567  
QY 216 -----VPSVESKTKPFTLPILTISELINSFPPIP-----IEQL 249  
Db 568 TTHISQSALPOGIEDLTSEVAQALTLIS-LPKQDSLPDITKASGPAHSKEVPALTAVETG 626  
QY 250 YT---APNET---NVQOQN-----GRCT---LDGE----- 271  
Db 627 ATNPLAPSDTVQTRHVVQRRSRSESTIESFFARGACVAIEVDNEQPTTTRAKLFAWMRI 686  
QY 272 -LQGTTLSSAVCFLOOR-----TVADNGNDWDLQLQTY-PNGASYDPT 316  
Db 687 TYKDTVQLARKLEFFETYSRDFMEFTFVVTAFTNANNHALNQ-VYQIMYIPPGA----- 740  
QY 317 DEVPAPLGTQDFSGMLYGVLTQDNVNVSTGBAKNAKGIYISTTSKFTPKIGSLHSIT 376  
Db 741 ---PTPKSWDDY-----TWQTSNPSIFITYGAAPARISVPVYGLANAY 781  
QY 377 EHVHPNQSRFTFVGAVDEN 397  
Db 782 SHFYDG---FAKVPLKTDAN 798

RESULT 11  
5516630-2  
;PATENT NO. 5516630  
;APPLICANT: TICEHURST, JOHN R.;BALTIMORE, DAVID;FEINSTONE,  
;STEPHEN M.;PURCELL, ROBERT H.;RACANIETTO, VINCENT R.;  
;BAROUDY, BAHIGE M.  
;TITLE OF INVENTION: METHODS OF DETECTING HEPATITIS A VIRUS  
;NUMBER OF SEQUENCES: 4  
;CURRENT APPLICATION DATA:  
;APPLICATION NUMBER: US/07/788,262  
;FILING DATE: 06-NOV-1991  
;PRIOR APPLICATION DATA:  
;APPLICATION NUMBER: 256,135  
;FILING DATE: 06-OCT-1988  
;APPLICATION NUMBER: 654,942  
;FILING DATE: 27-SEP-1984  
;APPLICATION NUMBER: 537,911  
;FILING DATE: 30-SEP-1983  
;SEQ ID NO.2:  
;LENGTH: 1091  
5516630-2

Query Match 3.8%; Score 108.5; DB 6; Length 1091;  
Best Local Similarity 21.6%; Pred. No. 0.31; 136; Indels 133; Gaps 22;  
Matches 87; Conservative 46; Mismatches 46; Mismatches 133; Gaps 22;

QY 96 YNGYAG-GVEVQVLLAGNAFTAGKILFAAIPPNFLVDMISPAQITMLPH-LIVDVRTLE 152  
Db 360 YHTYARFGIEIQVINPTFPQGGGLICAMVFGDSYSGI--ASLTVPYHGLLNCNINNV 417  
QY 153 PIMTLPDVRNVYHFNNQOPRMLVAMLYTPLRNSGDDVFTVSCRVLTRPTDPDEF 212  
Db 418 RIKVFFIYTRGA-YHFKDPQYPMVELTIRVWSEL-NIGTGTSAYT-SUNVLART-DUEL 473  
QY 213 IIVLPSPVESKTKP-----TLPILTISELTNSR----- 241  
Db 474 HGLTLPSTOMNEFRVSTTENVNLSNVEBARAKMSFALQOEDKWSQGGGIKITHF 533  
QY 242 -----FP-----IPIEQLYTAPNETVNVQCONCRCLTDELQGT 276  
Db 534 TTWTSIPTLAAQFPFNASDSVQGIKIVPVDPPYFQMTNTNPDQ-----KC-----IT 581  
QY 277 QLLS---SACVFCIQRTRVADNGNDWDLQLQTYPGA---SYDPTDEV----- 319  
Db 582 ALASICQMFCEWRGDLV-----FDQVFPFTKHSRLLFCFVPGNELIDVSGITLKQAT 635  
QY 320 PAPLGTQDFSGMLYGVLTQDNVNVSTGBAKNAKGIYISTTSKFTPKIGSLHSIT 363  
Db 636 TAPCAVMDITGVQSTLRFRVPMISDTPYRVNRYTKSA-HOKGEY--TAIGKLIVCYNRL 692

Db 741 ---PTPKSWDDY-----TWQTSNPSIFITYGAAPARISVPVYGLANAY 781  
QY 377 EHVHPNQSRFTFVGAVDEN 397  
Db 782 SHFYDG---FAKVPLKTDAN 798

RESULT 10  
US-08-465-250-2  
; Sequence 2, Application US/08465250  
; Patent No. 6136570  
; GENERAL INFORMATION:  
; APPLICANT: Racaniello, Vincent  
; APPLICANT: Tatem, Joanne M.  
; TITLE OF INVENTION: METHODS FOR PRODUCING RNA VIRUSES FROM  
; TITLE OF INVENTION: CDNA  
; NUMBER OF SEQUENCES: 9  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Cooper & Dunham LLP  
; STREET: 1185 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent in Release 1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,250  
; FILING DATE: 6-JUN-1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: White, John P.  
; REGISTRATION NUMBER: 28,678  
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 278-0400  
; TELEFAX: (212) 391-0525  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2206 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-465-250-2

Query Match 4.0%; Score 115; DB 3; Length 2206;  
Best Local Similarity 18.7%; Pred. No. 0.23;  
Matches 105; Conservative 65; Mismatches 179; Indels 212; Gaps 25;

QY 3 MASNDAAASNDGAASLVEGINEETPLPVPAGASIAAPVA-----GQTNIIIDP- 50  
Db 284 LAIDSMVXHNWGIALLP-----LSPLDPAQDSVSEIPIVTIAPMCSEFNGLRNVTA 338  
QY 51 -----WIRTNFVQAPNG--EFTVSPR-NSPGEI----- 75  
Db 339 FQGLPVLNTPGNSQVLTSDNHQSPCAIPEFDVTPPIDIPGEYKNNMELAEIDTWIPLNLE 398  
QY 76 -----LLNLELGPDLNPLYLAH-----LSRMVYAGVGEVQV 107  
Db 399 STKRNTMDMYRVLTSSADLSQIPILCLSLSPAFDRLSHTMLGEVLNYTHWAGSLKFTF 458  
QY 108 LLAGNAFTAGKILFAAIPPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVYH 167  
Db 459 LFCGSMWATGKILVAVAPPG-AQPPTSRKEAMLGTHVMDLGLQSSCTMVVFWISNVYR 517  
QY 168 FNNQPO-----PRRLVAMLYTPLRNSGDDVFTVSCRVLTRPTDPDEFIYL-- 215

<b>Qy</b>	364	-TPKIGSIGLHSITEHVHPN-----QQSRFTPVGVAVDENT	398
	:	:   :   :	:   :
<b>D<sub>b</sub></b>	693	TSPS-----NVASHVRVNVYLSAINLECFAPLYHAMDVTT	727

RESULT 12  
US-08-475-886-2  
; Sequence 2, Application US/08475896A  
; Patent No. 6113912  
; GENERAL INFORMATION:  
; APPLICANT: FUNKHOUSER, ANN W  
; APPLICANT: EMERSON, SUZANNE U  
; APPLICANT: PURCELL, ROBERT H  
; APPLICANT: D'HONDT, ERIC  
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

RESULT 13  
US-08-475-886-4  
; Sequence 4, Application US/08475886A  
; Patent No. 6113912  
; GENERAL INFORMATION:  
; APPLICANT: FUNKHOUSER, ANN W  
; APPLICANT: EMERSON, SUZANNE U  
; APPLICANT: PORCELL, ROBERT H  
; APPLICANT: D'HONDT, ERIC  
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES

```

; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475,886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947,338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397,232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 2227
; TYPE: PR1
; ORGANISM: Attenuated HAV (Pass 35), strain HM-175
; US-08-475-886-4

```

```

RESULT 14
US-08-475-886-6
; Sequence 6, Application US/08475886A
; Patent No. 6113912
; GENERAL INFORMATION:
; APPLICANT: FUNKHOUSER, ANN W
; APPLICANT: EMERSON, SUZANNE U
; APPLICANT: PURCELL, ROBERT H
; APPLICANT: D'HONDT, ERIC
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES
; FILE REFERENCE: 20264262US2
; CURRENT APPLICATION NUMBER: US/08/475, 886A
; CURRENT FILING DATE: 1995-06-07
; EARLIER APPLICATION NUMBER: 07/947, 338
; EARLIER FILING DATE: 1992-09-18
; EARLIER APPLICATION NUMBER: 08/397, 232
; EARLIER FILING DATE: 1995-03-10
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 2227
; TYPE: PRP
; ORGANISM: Attenuated (4380) HAV, strain HM-175
US-08-475-886-6

```

Query Match 3.8%; Score 108.5; DB 3; Length 2227;  
Best Local Similarity 21.6%; Pred. No. 1;  
Matches 87; Conservative 46; Mismatches 136; Indels 133; Gaps 22;  
QY 96 YNGYAG-GVEVQVLLAGNAFTAGKILFAAIPNPNFLVDMISPAQITMLPH--LIVDVRTLE 152  
DB 123 YHTYARFGIEIQVQINPTFPQGGGLICAMVFGDQSYGSI--ASLTVPYHGLLNCNINNV 180  
QY 153 PIMTLPDVRNVYHFNNQOPRMLVAMLYTPRLSNGSGDDVFTVSCRVLTRPTDFEF 212  
DB 181 RIKVPFIYTRGA-YHFKDPQYVWMLTIRVWSEL-NIGTGTSAYT-SLNVLARFT-DLEL 236  
QY 213 IYLVPPSVESKTKPF-----TLPILTISELTNSR----- 241  
DB 237 HGLTLPSTQMRNEFRVSTTENVNLSNVEDARAKMSFALDQEDWKSPPSQGGGIKITHF 296  
QY 242 -----FP-----IPLEQLYTAPNETNVVQCONGRCTLDGELQGT 276  
DB 297 TWTISPTLAAQFPFNASDSVGQIQKIVDPYFFQMTNTPDQ-----KC-----IT 344  
QY 277 QLLS--SAVCFLOGRTVADNGDNWQNLQLTYPNGA--SYDPTDEV----- 319  
DB 345 ALASICQMFCEWGRDLV-----FDQVFPPTKYHSGRLLFCFVPGNELIDVSGITLKQAT 398  
QY 320 PAPLGTQDPSGMLYGVL-----TQDNVNVSTGEAKNAKGIYISTTSKGF----- 363  
DB 399 TAPCAVMDITGVOSTLRFVRPWISDTPYRVNRYTKSA-HQKGEY--TAIGKLIYVCYNRL 455  
QY 364 -TPKIGSIGLHSITEHVHPN-----QOSRFTPVGVAVDENT 398  
DB 456 TSFS-----NVASHVRNVVYLSAINLECFAPLYHAMDVTT 490

Search completed: June 1, 2004, 13:58:06  
Job time : 14.9134 secs

RESULT 15  
US-08-397-232-2  
; Sequence 2, Application US/08397232A  
; Patent No. 6180110  
; GENERAL INFORMATION:  
; APPLICANT: FUNKHOUSER, ANN W  
; APPLICANT: EMERSON, SUZANNE U  
; APPLICANT: PURCELL, ROBERT H  
; APPLICANT: D'HONDT, ERIC  
; TITLE OF INVENTION: HEPATITIS A VIRUS VACCINES  
; FILE REFERENCE: 202642620S1  
; CURRENT APPLICATION NUMBER: US/08/397,232A  
; CURRENT FILING DATE: 1995-04-17  
; EARLIER APPLICATION NUMBER: 07/947,338  
; EARLIER FILING DATE: 1992-09-18  
; EARLIER APPLICATION NUMBER: PCT/US93/08610  
; EARLIER FILING DATE: 1993-09-17  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 2227  
; TYPE: PRT  
; ORGANISM: WILD-TYPE HUMAN HEPATITIS A VIRUS, STRAIN HM-175  
US-08-397-232-2  
Query Match 3.8%; Score 108.5; DB 3; Length 2227;  
Best Local Similarity 21.6%; Pred. No. 1;  
Matches 87; Conservative 46; Mismatches 136; Indels 133; Gaps 22;  
QY 96 YNGYAG-GVEVQVLLAGNAFTAGKILFAAIPNPNFLVDMISPAQITMLPH--LIVDVRTLE 152  
DB 123 YHTYARFGIEIQVQINPTFPQGGGLICAMVFGDQSYGSI--ASLTVPYHGLLNCNINNV 180  
QY 153 PIMTLPDVRNVYHFNNQOPRMLVAMLYTPRLSNGSGDDVFTVSCRVLTRPTDFEF 212  
DB 181 RIKVPFIYTRGA-YHFKDPQYVWMLTIRVWSEL-NIGTGTSAYT-SLNVLARFT-DLEL 236  
QY 213 IYLVPPSVESKTKPF-----TLPILTISELTNSR----- 241  
DB 237 HGLTLPSTQMRNEFRVSTTENVNLSNVEDARAKMSFALDQEDWKSPPSQGGGIKITHF 296

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.648 Seconds  
(without alignments)  
4368.312 Million cell updates/sec

Title: US-09-926-799-9  
Perfect score: 2863  
Sequence: 1 MKMASNDAPSDGAASLVP.....VNQFSLAPMTGNGRRRIQ 542

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:\*

1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1179	41.2	530	14	US-10-314-739-3
2	276.5	9.7	547	12	Sequence 3, Appli
3	276.5	9.7	671	12	Sequence 4, Appli
4	265.5	9.3	668	14	Sequence 2, Appli
5	265.5	9.3	669	14	Sequence 4, Appli
6	145.5	5.1	1147	15	Sequence 2, Appli
7	133	4.6	6310	12	Sequence 38, Appli
8	116	4.1	394	15	Sequence 6793, A
9	114.5	4.0	594	12	Sequence 3560, Ap
10	112.5	3.9	3930	12	Sequence 65399, A
11	111.5	3.9	3317	15	Sequence 46817, A
12	110.5	3.9	5376	15	Sequence 200, App
13	110.5	3.9	5376	15	Sequence 74, Appli
14	110.5	3.9	6304	14	Sequence 74, Appli
15	109.5	3.8	1109	16	Sequence 16, Appli
					Sequence 1196, Ap

16 109.5 3.8 1109 16 US-10-389-566-2378 Sequence 2378, Ap  
17 109.5 3.8 2468 12 US-10-282-122A-66335 Sequence 66335, A  
18 109.5 3.8 2468 14 US-10-246-330-4 Sequence 4, Appli  
19 108.5 3.8 836 14 US-10-272-459-40 Sequence 40, Appli  
20 108.5 3.8 980 14 US-10-272-459-41 Sequence 41, Appli  
21 108.5 3.8 2227 9 US-09-929-955-12 Sequence 12, Appli  
22 108.5 3.8 2227 13 US-10-104-366-12 Sequence 12, Appli  
23 108.5 3.8 2227 13 US-10-135-988-2 Sequence 2, Appli  
24 108.5 3.8 2227 13 US-10-135-988-4 Sequence 4, Appli  
25 108.5 3.8 2227 13 US-10-135-988-6 Sequence 6, Appli  
26 108.5 3.8 2227 16 US-10-719-619-12 Sequence 12, Appli  
27 107.5 3.8 2777 12 US-10-423-483-4 Sequence 4, Appli  
28 107.5 3.8 2777 15 US-10-220-587-4 Sequence 4, Appli  
29 107.5 3.8 3399 15 US-10-080-334-196 Sequence 196, App  
30 105.5 3.7 368 14 US-10-156-761-7936 Sequence 7936, Ap  
31 105 3.7 547 15 US-10-080-334-276 Sequence 276, App  
32 105 3.7 596 12 US-10-188-248-68 Sequence 68, Appli  
33 104.5 3.7 1194 12 US-10-282-122A-46577 Sequence 46577, A  
34 104 3.6 925 15 US-10-108-260A-3889 Sequence 3889, Ap  
35 103 3.6 624 15 US-10-258-106-4 Sequence 4, Appli  
36 103 3.6 913 15 US-10-369-493-2294 Sequence 2294, Ap  
37 102.5 3.6 698 15 US-10-074-978A-144 Sequence 144, App  
38 101.5 3.5 390 11 US-09-764-875-675 Sequence 675, App  
39 101.5 3.5 1469 12 US-10-262-511-164 Sequence 164, App  
40 101.5 3.5 2753 12 US-10-262-511-160 Sequence 160, App  
41 101.5 3.5 2753 12 US-10-262-511-166 Sequence 166, App  
42 101.5 3.5 3838 12 US-10-262-511-162 Sequence 162, App  
43 101.5 3.5 5374 15 US-10-028-248A-75 Sequence 75, Appli  
44 101.5 3.5 5374 15 US-10-107-782-75 Sequence 75, Appli  
45 101 3.5 1004 12 US-10-114-270-32 Sequence 32, Appli

#### ALIGNMENTS

RESULT 1  
US-10-314-739-3  
; Sequence 3, Application US/10314739  
; Publication No. US20030129588A1  
; GENERAL INFORMATION:  
; APPLICANT: Bates, Mary K  
; Jiang, Xi  
; Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; Characterize No. US20030129588A1walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pubbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE: <Unknown>  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/10/314,739  
; APPLICATION NUMBER: US/10/314,739  
; FILING DATE: 09-Dec-2002  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX: <Unknown>





Qy	11	SNDGAASLVPGINETMP-----LEPVAGASTAAPVAGOTNIIDPW-IRTNFVOAPNGEPT	65
Db	125	ADDGSIITAPEGQ---TWGGVGIAEPSQAQMSTAADMATGKSDVSEWAFPSFTSVNW---	178
Qy	66	VSPRNSPGEIILLNELGPDLPANPLYLAHLSRMVNGYAGGVEQVLLAGNAFTACKILFAAPT	125
Db	179	-STSETQGKIIFUKOSGLPPLNLYLHLAKLYAVWMSGIEVRFSISGSGVFGKLAAIVVP	237
Qy	126	PNFLVDMSIPAQITMLPHLIVDVRTLEPIMTPLDPVRNVFYHFNQOPPRMELVAMLYTP	185
Db	238	PG--VDPVQSTSMLOYPHVLFDAQRVEPIFCLPDLRLTYLHMSDTD-TTSLVIMVYND	294
Qy	186	LRSNGSGDDVTVC--RVLTRPTDPDFRYLVLP-----SVGS-----KT-----	224
Db	295	L-INPYANDANSGCCIVTVETKPGPDFHLLKPPGSMLTGHSIPSDLIPTKTSMLIGNR	353
Qy	225	-----KPFTL-----PI-LTISELTNSRFPIPIEQLYTA	252
Db	354	YWSDIDEVIRPFVFQANRHFDPNQETAGWSTPRFRPISVLITEONGAKLGIVATDYIV	413
Qy	253	PN-----ETNVQCQNGRCTLDGEL-----OGTTQLSSAFCFLQGRTVAD---N	294
Db	414	PGIPDGWPDT-----TIPGELIPAGDYAITNGTGNDITTAT----GYDTADIKN	459
Qy	295	GDN-----W-DONLLQLUTYPNGASYD-PTDEVPAPLGTQDFSGMLYGVLTDQ	339
Db	460	MNTNFRGMVIGCSLORAWGDKKISNTAFITTATLDGDNNNKINPCNTIDQSKIT---VWFQD	516
Qy	340	N-----VNVS-----TGEAKNKG-----IYISTSGKFPKGISGLHSITEHV	379
Db	517	NHVCKKAQTSDDTLALLGYTGIGEAQISDRDRVRIST-----LPEFGARGN-----	565
Qy	380	HPNQOCSRFTPVGVAVENTPFPQQWLPHYAGSLALNTNLAP	420
Db	566	HPIEFKKSIKLGYVIRSIDVFNISOIL-HTSROLSLNHYLLP	605

RESULT 4  
 US-10-209-507-4  
 ; Sequence 4, Application US/10209507  
 ; Publication No. US20030105033A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Audonnet, et al.  
 ; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V  
 ; FILE REFERENCE: 454313-3151.2  
 ; CURRENT APPLICATION NUMBER: US/10/209,507  
 ; CURRENT FILING DATE: 2002-11-12  
 ; PRIOR APPLICATION NUMBER: 09/617,594  
 ; PRIOR FILING DATE: 2000-07-14  
 ; PRIOR APPLICATION NUMBER: 60/193,332  
 ; PRIOR FILING DATE: 2000-03-30  
 ; PRIOR APPLICATION NUMBER: France 00 01761  
 ; PRIOR FILING DATE: 2000-02-11  
 ; PRIOR APPLICATION NUMBER: France 99 09421  
 ; PRIOR FILING DATE: 1999-07-16  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 4  
 ; LENGTH: 668  
 ; TYPE: PRT  
 ; ORGANISM: Feline calicivirus  
 ; US-10-209-507-4

```

Db      181  SETGKILFKQSLGFLNPNYLTHLAKLVAVWSGSIEVRFSGSGVFGCKLAAIVVPPG- 239
Qy      129  LVDMSPAQITMLPHLIVDVRTLEPIMTPLPDVRNVFYHFNNQPQRMRLVAMLYTPL-- 186
Db      240  -IDPVQSTMLQYPHVLFDARQVPEVFTIEDLRNSLTHLMSDTD-TTSLVIMINYNDLN 297
Qy      187  ----RNSGDDVFTVSCRVLTRTPDPEFTYLVPPSVESKTKPFTLPILTTISELTSNRP 242
Db      298  PYANDSNSGCIV-----TVETKGPDPFKPHLLKPPG-----SMLTHGSI 337
Qy      243  P 243
Db      338  P 338

RESULT 5
US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RES
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
; US-10-209-507-2

```

374	HFDPNQETAGWSTPRFRPITITIIISSENGSKLGTGVATDYIVPG-----IPDGM	421
Db		
272	LQGTTLQLLSAVCFLOQRTVADNGDNQDLQLTYPNGA-----SYDPTDVPAPL	323
Qy		
422	PDTTI-----GEELTPAGD-----YSITNGSGNDIATANAYDSADYI---T	459
Db		
324	GTDPFGM-LYGVLTQDNVNVSTCEAKNAKGIYISTT--SGKFTPKISGISLHSTEHVH	380
Qy		
460	NTNFRGMYICGALOR-----AWGDKISSTAFITTAKEGNTKLKPSNTIDMTKI---	509
Db		

Wed Jun 2 09:13:34 2004

```

Db      500 ENDTILL---VSASAGKO-FEURL 518

RESULT 7
US-10-282-122A-67793
; Sequence 67793, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67793
; LENGTH: 6310
; TYPE: PRT
; ORGANISM: Pseudomonas putida
; US-10-282-122A-67793

Query Match      4.6%; Score 133; DB 12; Length 6310;
Best Local Similarity 23.2%; Pred. No. 0.13;
Matches 113; Conservative 58; Mismatches 201; Indels 116; Gaps 26;

QY      7 DAAPS-----NDGASLVEGINETMPEPVAGASIAAPVAGQTNIIDPWIRTNFVQAPN 61
Db      3575 NASPSATFTADSTAPAPAGLITIPGSAIOGSAGSTV-EVKLADGTLGVTVVPAG 3633
QY      62 GEFTV--SPRNSPGEILLNLEL-----GPDLPYLAHLRMYNGYAGGV 105
Db      3634 GSTVPLSPAQLDQG-ALNVTLTDAAGNISQPSQIFAPDITP-----PALPTDV 3681
QY      106 QVLLAGNAFT-----AGKILFAAIPP--NELVDMISPAQ-----ITMLPHLI 145
Db      3682 AVSSDGTAVTGNAPGASSVTVSDGAGNVITAVNPDGSFVPLDTPQNNQGVTV---VV 3738
QY      146 VDV--RTLEPIMTLPDVRNVFYHFNQPPQPRMLVAMLYTPLRS--NCGSDDVFTVSCR 201
Db      3739 TDAAGNDSAPFVSVTAPDTTN-----PEPATGLTV---SPDGSVTGGTAEPGSTVEVR 3787

US-10-327-481A-38
; Sequence 38, Application US/10327481A
; Publication No. US20040001864A1
; GENERAL INFORMATION:
; APPLICANT: King, Andrew M.O.
; APPLICANT: Burman, Alison J.
; APPLICANT: Audonnet, Jean-Christophe F.
; APPLICANT: Lombard, Michel F.A.
; TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease
; FILE REFERENCE: 454313-3178
; CURRENT APPLICATION NUMBER: US/10/327,481A
; CURRENT FILING DATE: 2002-12-20
; PRIOR APPLICATION NUMBER: PCT/FR01/02042
; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: FR 0008437
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 1147
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
; US-10-327-481A-38

Query Match      5.1%; Score 145.5; DB 15; Length 1147;
Best Local Similarity 20.8%; Pred. No. 0.00056;
Matches 92; Conservative 47; Mismatches 203; Indels 101; Gaps 13;

QY      70 NSPGEILLNLBELGPDLPYLHLRSMYNGYAGGVQVLLAGNAFTAGKILFAAIPNPL 129
Db      158 DKPFGYLTKELEPTDHHGVFGLHVDSDYAYMRNGMDVEVSAGVNGFCGLLVAMVPEWKA 217
QY      130 VDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPPQPRMLVAMLYTPLRSN 189
Db      218 FDTREKYQUTLFFHQFI-----SPRTNMTAHTVYPLGV 251
QY      190 GSGDDVFTVSCRVLRTPTDPEFIVLPSPVESKTKPTLPILITISELTNSRFPPIEQOL 249
Db      252 NRYDQY-----KKHKPTLVVMVLSPLUTSVNTAAPQIKV 285
QY      250 YT--APNETNVQCGNCRCTLDGELQGTTLQLLSAVCFELQRTVADNGDNWDQNLLOQTY 307
Db      286 YANIAPTYVHVA-----GELPSKEGIFPVACADYGGLVTTDPKTAADPVYGVKN 335
QY      308 PNGASY-----DPTDEVPAPLGTQDFSGMLYGVLTQDNV-----NVSTGEAKNAKG 353
Db      336 PPKTNPYGRFTNLLDVAEACPTFLREDD--GKPYVTVTRADDTRLLAKFDVSLA-AKHSN 392
QY      354 IYISTSGKFTPKIGSIGHLSI---TEHVHPNQSRFTVGVAVDENTPFQOWVLPHYAG 410
Db      393 TYLSGIAQYQYQYSGTINLHFMFTGTDSKARYWYAIIPGVETPDPPEEAHCIIHAEW 452
QY      411 SLALNTNLAPAVAPTEPGEQLLFFRSRVCVQSLQGDADIFDCLLPQEWNVHIFYEAPRS 470
Db      453 DTGLNSKFTESI--PYVSAADYATASDTAETTNVQ-----WVCVYQITHGKA 499
QY      471 QADVALLIRYNPDGRTLFEAKL 493

```



Wed Jun 2 09:13:34 2004

Db 290 RINLKFTLTLDMKTEEFKVRGAADMKGRVVALRVNPNVGNKSRKEIDYTKFVGIYA 349  
Qy 329 SCMLYGLVTQDNVNSTGE-----AKNAKIYISTTSKGF 363  
Db 350 KGLAY-IKVNDAGNLSGDSGLQSPIVKFLSENALKELIERTAAQNGDIIFFGADKAKV 408  
Qy 364 TPK-IGSIGLHSITHHVHPNQSRFTFVGVAVENTPPQWVL 405  
Db 409 VNEAVGALKRIKVGLEHGKDN-----GYFTDEWKPL--WV 441

RESULT 10  
US-10-282-122A-46817  
; Sequence 46817, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangau  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.0344  
; CURRENT APPLICATION NUMBER: US/10/282.122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 46817  
; LENGTH: 3930  
; TYPE: PRT  
; ORGANISM: Bacillus anthracis  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (19)..(19)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (38)..(38)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (3455)..(3455)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE

; LOCATION: (3459)..(3459)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (3552)..(3552)  
; OTHER INFORMATION: X=any amino acid  
; FEATURE:  
; NAME/KEY: MISC FEATURE  
; LOCATION: (3566)..(3566)  
; OTHER INFORMATION: X=any amino acid  
US-10-282-122A-46817  
Query Match 3.9%; Score 112.5; DB 12; Length 3930;  
Best Local Similarity 18.9%; Pred. No. 5.4;  
Matches 136; Conservative 91; Mismatches 262; Indels 229; Gaps 34;  
Qy 9 APSNDGAAS-----LVPEG---INETMPL-----EPVAGASTAAAPVAGOT---- 45  
Db 2203 ALTNNNGSSAQNVIPTDTVPSTAFIADTFSGINGIPQSGAMPVNGVNGISITAGTTTVS 2262  
Qy 46 -----NLIIDPWIRTF--VOAPNGEFTVSPRNSP--GEILLNLELGPDLN---- 86  
Db 2263 FQVTVTSLPTENPIVNFSSISYQLVSPDPAETISINPVSTQIKAILSMKNESLSFANI 2322  
Qy 87 ----PYLAHLRMYNGYAGG-VEVOVLLAGNAFTAGK-----ILFAAIP 125  
Db 2323 GQTAFYTTTISINIGNTDATNIVFTDVLPLNGVTFVPNTLTVGVLPDANPNTGVLLATLP 2382  
Qy 126 PNFVLDMISPAQITMLPHLIVDVRTLEPIMTPLPDVNRNVYHFNQSP--QPRMLVAMLY 183  
Db 2383 PNEYISIV--FQVT-----VNSIPPI-NEAPNTASTTYEFTVDVPPVPPVLSAATSNT 2431  
Qy 184 TPLRSNGSGDDVFTVSCRVLTRTPDFEF-----NANIISTKTDLTADVGNTTFTLNPNTGNVTATDVTVIDTLD 2482  
Db 2432 TLLQIN-----NANIISTKTDLTADVGNTTFTLNPNTGNVTATDVTVIDTLD 2482  
Qy 213 --IYLVPPS-----VESKTPFLPI---LTISELT 238  
Db 2483 SNLTFVPNSFTVNGQITLNADLSTGVNIGSINGGTAAIVTFQATVTLPLNPFISNSALT 2542  
Qy 239 NSRPPI-PIEQLYTAPNETNVVQCNGRCITLDGELQGTOLLSSAVCFLOGRVADNGD- 296  
Db 2543 TYRIVDPDQSPITTSNGSNTTTQINSAILTAQKSTNVFTVDIGDIVYSVTITNSGV 2602  
Qy 297 NWDQNLQLTYPNGASYDP-----TDEV 319  
Db 2603 NATNVIFTDVIPTDGTSTFEPNSFTLNGTIENANITGVPIGDIAPNESAIVEPHITSNEI 2662  
Qy 320 PA--PLGTQDFSGMLYGVLTQDNVNVSTGEAKNAGIYISTTSCK---FTPKIGSIGLHS 374  
Db 2663 PAINFI-TNQASVSFQHVNPANPPVS-----KNITSNSVTITTESAILTTTKIGDKAFAT 2717  
Qy 375 ITEVHPNQSRFTFVGVAVENT-----PFQWVLPVHVGSLAENTNLAPAVAPT----- 425  
Db 2718 IGDTI--TVTTITNIGNIPANNVIFSDPIPSWT-QFVAGSVYVDGTPLPSASITSGIGI 2774  
Qy 426 ----FPGEQL-LFERSRVPVQVQGLQGDADFIDCLLPQEWVNHFYQ-----EAPSQADVAL 476  
Db 2775 NTIIPNQTVTIIFQ-----QVIVSNPPTFPELQNLAFVNFQYVNGVALQAPQGNVETNV 2829  
Qy 477 IRYVNPDTGRTLFEAKLHRSFGITVSHTCAYPLVPPNCHFRFDSWVNOFYSLAPMGT 534  
Db 2830 --FVTAIHSALLSAVKTAFASTAFANIGDITTYTVLQNSGNTNATN-VN-FSDLIFFGT 2883

RESULT 11  
US-10-080-334-200  
; Sequence 200, Application US/10080334  
; Publication No. US20040002584A1  
; GENERAL INFORMATION:  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shinkets, Richard A  
; APPLICANT: Li, Li

APPLICANT: Shenoy, Suresh G  
APPLICANT: Kekuda, Ramesh  
APPLICANT: Spytek, Kimberly A  
APPLICANT: Vernet, Corine A. M.  
APPLICANT: Malyankar, Uriel M  
APPLICANT: Guo, Xiaojia  
APPLICANT: Gusev, Vladimir Y  
APPLICANT: Casman, Stacie J  
APPLICANT: Boldog, Ferenc L  
APPLICANT: Furtak, Katarzyna  
APPLICANT: Tchernev, Velizar T  
APPLICANT: Patturajan, Meera  
APPLICANT: Gangolli, Esha A  
APPLICANT: Padigaru, Muralidhara  
APPLICANT: Liu, Xiaohong  
APPLICANT: Baumgartner, Jason C.  
APPLICANT: Gerlach, Valerie  
APPLICANT: Spaderna, Steven K  
APPLICANT: Zerhusen, Bryan D  
TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of  
FILE REFERENCE: 21402-275  
CURRENT APPLICATION NUMBER: US/10/080,334  
PRIOR FILING DATE: 2002-02-21  
PRIOR APPLICATION NUMBER: 60/270,523  
PRIOR FILING DATE: 2001-02-21  
PRIOR APPLICATION NUMBER: 60/322,712  
PRIOR FILING DATE: 2001-09-17  
PRIOR APPLICATION NUMBER: 60/311,980  
PRIOR FILING DATE: 2001-08-13  
PRIOR APPLICATION NUMBER: 60/330,307  
PRIOR FILING DATE: 2001-10-18  
PRIOR APPLICATION NUMBER: 60/278,796  
PRIOR FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: 60/281,521  
PRIOR FILING DATE: 2001-04-04  
PRIOR APPLICATION NUMBER: 60/276,677  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 60/311,595  
PRIOR FILING DATE: 2001-08-10  
PRIOR APPLICATION NUMBER: 60/270,220  
PRIOR FILING DATE: 2001-02-21  
PRIOR APPLICATION NUMBER: 60/274,295  
PRIOR FILING DATE: 2001-03-08  
PRIOR APPLICATION NUMBER: 60/318,526  
PRIOR FILING DATE: 2001-09-10  
PRIOR APPLICATION NUMBER: 60/286,548  
PRIOR FILING DATE: 2001-04-25  
PRIOR APPLICATION NUMBER: 60/291,765  
PRIOR FILING DATE: 2001-05-17  
PRIOR APPLICATION NUMBER: 60/270,797  
PRIOR FILING DATE: 2001-02-23  
PRIOR APPLICATION NUMBER: 60/276,400  
PRIOR FILING DATE: 2001-03-16  
PRIOR APPLICATION NUMBER: 60/270,810  
PRIOR FILING DATE: 2001-02-23  
NUMBER OF SEQ ID NOS: 388  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 200  
LENGTH: 3317  
TYPE: PRT  
ORGANISM: Rattus norvegicus  
US-10-080-334-200

Query Match 3.9%; Score 111.5; DB 15; Length 3317;  
Best Local Similarity 19.2%; Pred. No. 5.1;  
Matches 126; Conservative 91; Mismatches 227; Indels 213; Gaps 35;

QY 6 NDAAPSDGAAALVPEGINETWPLPEPVAG--ASIAAPVAGQTNIIDPWIRTNFVQAPNGE 63  
DB 43 SEDTPGVSSVTQLLRDMDNDPLVGVSGEASRFAPVPTGVV--WLRQPLDRETKSE 100  
QY 64 FTV--SPRNSPGEIL--LNLELGPDLN-----PYLAHLRMYNGYAGGVEVQVLLA 110

DB 101 FTVEFSVDHOGVITRKVNIQVG--DVNDNAPTFFHNQPSV---RIPENTPVGTPIFVNA 156  
QY 111 --GNAFTAGKILFAAIPPN--FLVD-----MISPAQITM----- 140  
DB 157 TDPDLGAGGSVLVYSPQSPQFAIDSARGIVTVIRELDYEVYTAQVLTNATDQDKTRPL 216  
QY 141 -----LPHLIVDVRTLEPMTPLPDVRNVFVHNNQPPRMRLVAML-----YT 184  
DB 217 STLANLAIITDQMDPIFINLPYSTNIYEH--SPPGTTVRVITAVDQDKGRPRGIGYT 274  
QY 185 PLRNSGSGDDVFTVSC-----RVLTPTPDPE--FIYLV-----PPSVESKTK 225  
DB 275 IV--SGNTNSIFALDYISGALTNLGLDRENPLVSHGHFILTVMKTELDNDRSPSDATVTT 332  
QY 226 PFTLPILTIS---ELTNSRFPPIEQI-----YTAPNETNVVQCNGRCTL----- 268  
DB 333 TFNILVIDINDNAPEFNSSEYSVAITELAQVGFALPLFIQVVVDKDEGLNSMEFVYLVGN 392  
QY 269 -----DGELOQTQLLSSAVCFLOGRVTADNGNDQNL--QLTYPNGASY----- 313  
DB 393 SHHFIISPTSVQKADIRIRVAIPLDYETV---DRYDFDLFANESVDPDHVGYAKKITL 448  
QY 314 -DPTDEVPAPLGTQDFSCMLYGLVLTQDNVNVST-----GEAKNAKGIYS 357  
DB 449 INENDNRPI-----FSQPLYNVSLYENITVGTSLVTLATDNDVGTTEGV-----NYPPS 498  
QY 358 TTSKGTPTKIGSIGHSTIEHVHPNQQRFPVGVAVD---ENTPQQWVLPHYAGSLAL 414  
DB 499 DDPDRFSLD-KDTGLIMLIARLDYELIQRFILTVIARDGGGEET-----TGRVRI 547  
QY 415 MTNLAPAVATFPGEQLLFFRSVPCVQGLQGQDAFIDCLLPQEWNNHFYQGAAPSQADV 474  
DB 548 N-----VLDVNDNVPTFQ---KDAYVGAL-----RENEPSVTQL 578  
QY 475 ALIR-----YUNPDTGRTLFEAKLHRS-GFITVSHTGAYPLVVPNG 515  
DB 579 VRLRATDESDPPNNLITYSVNASAFGSYFDISVYEGYVISVSRPLDYEQI--PNG 633

## RESULT 12

US-10-028-248A-74

; Sequence 74, Application US/10028248A

; Publication No. US20030235882A1

; GENERAL INFORMATION:

; APPLICANT: Shimkete, Richard

; APPLICANT: Patturajan, Meera

; APPLICANT: Vernet, Corine

; APPLICANT: Casman, Stacie

; APPLICANT: Malyankar, Uriel

; APPLICANT: Shenoy, Suresh

; APPLICANT: Spytek, Kimberly

; APPLICANT: Gangolli, Esha

; APPLICANT: Miller, Charles

; APPLICANT: Boldog, Ferenc

; APPLICANT: Li, Li

; APPLICANT: Taupier Jr, Raymond J

; APPLICANT: Kekuda, Ramesh

; APPLICANT: Smithson, Glenda

; APPLICANT: Zerhusen, Bryan

; APPLICANT: Liu, Xiaohong

; APPLICANT: Colman, Steven

; APPLICANT: Tchernev, Velizar

; APPLICANT: Si, Jingsheng

; APPLICANT: Edinger, Shlomit

; APPLICANT: Stone, David

; APPLICANT: Sciore, Paul

; APPLICANT: Millet, Isabelle

; APPLICANT: Rothenberg, Mark

; TITLE OF INVENTION: No. US20030235882A1el Nucleic Acids and Polypeptides and Methods

; TITLE OF INVENTION: Theroof

; FILE REFERENCE: 21402-222

; CURRENT APPLICATION NUMBER: US/10/028,248A

Wed Jun 2 09:13:34 2004

```

; CURRENT FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311266
; PRIOR FILING DATE: 2001-08-09
; NUMBER OF SEQ ID NOS: 211
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 74
; LENGTH: 5376
; TYPE: PRT
; ORGANISM: Mus musculus
US-10-028-248A-74

Query Match 3.9%; Score 110.5; DB 15; Length 5376;
Best Local Similarity 20.9%; Pred. No. 14;
Matches 91; Conservative 56; Mismatches 177; Indels 111; Gaps 20;

QY 26 TUPLE-PVAGASIAAPVAGQNIID--PWIRNFVQAPNGEFTVSPRNS--PGEILLNLE 80
DB 683 TVTPEVPIVLEATAFPTGETTLYTEVPTVPTVGVHTEVTVNSPEETSVPTEETISTE 742
QY 81 LGPDLPYLAHLRMYNGYAGGVEQ-VLLAGNAFTAGKI-LFAAIPP-----NFLV 130
DB 743 V-TTVSPEETVP-----TEVPIVLEATASPTGEITLYTEVPTVGVHTEV 792
QY 131 DMISPAQITMLP---HLIVDVRTLEPIMTLP-----DVRNVFYHFNNOQPPRMLV 179
DB 793 TNSVPEE-TSVPEETISTEVTTSPEETTLPTPEVTVSTEVTVNSPEETSVP-PEETIL 850
QY 180 AMLYT-----PLRSGSGDDVFTVSCVLTPTPDPEFIYLVPPSVESKTKPTLPILTI 234
DB 851 TTLTYTEVPTVPTVGVHTEVTVNSPEETSVPTEETISTEVTTSPEETTLPTPEVTVS- 909
QY 235 SELTNSRFPPIEOLYAPNETNVVQCONGRCTLDGELQGTQLLSSAVCFLOQRTVADN 294
DB 910 TEVTN-----VSPETSVPPEETILTEI-----TTVSPEETVPPIEGTTL--- 949
QY 295 GDNWDQNLQLTYPNGASYDPTDEVAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGI 354
DB 950 -----PTEVLTVPIEVTTF-----PTGE----- 967
QY 355 YISTTSKFTPKIGS--IGLHSITEHVHPNQSRFTPVGVAVDENTFPQOWVLPVHAGSL 412
DB 968 ---TTVPTTEVPTVSTMTGVHTEVTVTFPEETSIPTEVATVLPASIPPEETTTTTEVTTT 1024

QY 413 ALNTNLAPAVATFP 427
DB 1025 PPEETIPEVTTVP 1039

RESULT 13
US-10-107-782-74
; Sequence 74, Application US/10107782
; Publication No. US20040018970A1
; GENERAL INFORMATION:
; APPLICANT: Boldog, Ferenc,
; APPLICANT: Casman, Stacie
; APPLICANT: Colman, Steve,
; APPLICANT: Edinger, Shlomit,
; APPLICANT: Gangolli, Esha,
; APPLICANT: Kekuda, Ramesh,
; APPLICANT: Li, Li,
; APPLICANT: Liu, Xiaohong,
; APPLICANT: Malyankar, Uriel,
; APPLICANT: Miller, Charles,
```

```

; APPLICANT: Millet, Isabelle,
; APPLICANT: Patturajan, Meera,
; APPLICANT: Rothenberg, Mark,
; APPLICANT: Sciore, Paul,
; APPLICANT: Sheno, Suresh,
; APPLICANT: Shimkets, Richard,
; APPLICANT: Si, Jingsheng,
; APPLICANT: Smithson, Glenda,
; APPLICANT: Spytek, Kimberly,
; APPLICANT: Stone, David,
; APPLICANT: Taupier, Raymond, Jr.,
; APPLICANT: Tchernev, Velizar,
; APPLICANT: Vernet, Corine,
; APPLICANT: Zerhusen, Brian
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES AND METHODS OF USE THEREOF
; FILE REFERENCE: 21402-222CIP
; CURRENT APPLICATION NUMBER: US/10/107,782
; CURRENT FILING DATE: 2002-03-27 248
; PRIOR APPLICATION NUMBER: 10/028,248
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 60/256,619
; PRIOR FILING DATE: 2000-12-19
; PRIOR APPLICATION NUMBER: 60/262,959
; PRIOR FILING DATE: 2001-01-19
; PRIOR APPLICATION NUMBER: 60/272,408
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: 60/285,189
; PRIOR FILING DATE: 2001-04-20
; PRIOR APPLICATION NUMBER: 60/308,039
; PRIOR FILING DATE: 2001-07-26
; PRIOR APPLICATION NUMBER: 60/311,266
; PRIOR FILING DATE: 2001-08-09
; PRIOR APPLICATION NUMBER: 60/279,344
; PRIOR FILING DATE: 2001-03-28
; NUMBER OF SEQ ID NOS: 215
; SOFTWARE: Curaseqlist version 0.1
; SEQ ID NO 74
; LENGTH: 5376
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-107-782-74

Query Match 3.9%; Score 110.5; DB 15; Length 5376;
Best Local Similarity 20.9%; Pred. No. 14;
Matches 91; Conservative 56; Mismatches 177; Indels 111; Gaps 20;

QY 26 TUPLE-PVAGASIAAPVAGQNIID--PWIRNFVQAPNGEFTVSPRNS--PGEILLNLE 80
DB 683 TVTPEVPIVLEATAFPTGETTLYTEVPTVPTVGVHTEVTVNSPEETSVPTEETISTE 742
QY 81 LGPDLPYLAHLRMYNGYAGGVEQ-VLLAGNAFTAGKI-LFAAIPP-----NFLV 130
DB 743 V-TTVSPEETVP-----TEVPIVLEATASPTGEITLYTEVPTVGVHTEV 792
QY 131 DMISPAQITMLP---HLIVDVRTLEPIMTLP-----DVRNVFYHFNNOQPPRMLV 179
DB 793 TNSVPEE-TSVPEETISTEVTTSPEETTLPTPEVTVSTEVTVNSPEETSVP-PEETIL 850
QY 180 AMLYT-----PLRSGSGDDVFTVSCVLTPTPDPEFIYLVPPSVESKTKPTLPILTI 234
DB 851 TTLTYTEVPTVPTVGVHTEVTVNSPEETSVPTEETISTEVTTSPEETTLPTPEVTVS- 909
QY 235 SELTNSRFPPIEOLYAPNETNVVQCONGRCTLDGELQGTQLLSSAVCFLOQRTVADN 294
DB 910 TEVTN-----VSPETSVPPEETILTEI-----TTVSPEETVPPIEGTTL--- 949
QY 295 GDNWDQNLQLTYPNGASYDPTDEVAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGI 354
DB 950 -----PTEVLTVPIEVTTF-----PTGE----- 967
QY 355 YISTTSKFTPKIGS--IGLHSITEHVHPNQSRFTPVGVAVDENTFPQOWVLPVHAGSL 412
DB 968 ---TTVPTTEVPTVSTMTGVHTEVTVTFPEETSIPTEVATVLPASIPPEETTTTTEVTTT 1024
```

QY 413 ALNTNLAPAVATPP 427  
Db 1025 PPEETTPAEVTTVP 1039

RESULT 14  
US-10-147-026-16  
; Sequence 16, Application US/10147026  
; Publication No. US20030003538A1  
; GENERAL INFORMATION:  
; APPLICANT: Dietrich, Paul Shartzner  
; APPLICANT: Huang, Chiao-Chain  
; APPLICANT: Johnson, Carl D.  
; APPLICANT: Sangameswaran, Lakshim  
; TITLE OF INVENTION: Neuropathic Pain Genes; Compositions  
; FILE REFERENCE: Thereof; Related Reagents  
; CURRENT APPLICATION NUMBER: US/10/147,026  
; PRIORITY FILING DATE: 2002-05-15  
; PRIOR APPLICATION NUMBER: 60/155,702  
; PRIORITY FILING DATE: 1999-09-23  
; PRIOR APPLICATION NUMBER: 60/189,931  
; PRIORITY FILING DATE: 2000-04-04  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16  
; LENGTH: 6304  
; TYPE: PRT  
; ORGANISM: human  
US-10-147-026-16

Query Match 3.9%; Score 110.5; DB 14; Length 6304;  
Best Local Similarity 19.5%; Pred. No. 18;  
Matches 104; Conservative 75; Mismatches 179; Indels 175; Gaps 25;  
QY 64 FTVSPRNS-----PGE-----ILLNLELGPDLNPNYLAHLRMYNGVAGVE 104  
Db 811 YRVEPRDSNEFYGNTGVLEFKPGEREIVITLLARLDGIPELDEH----- 854  
QY 105 VQVLNAGNATAGKILFAAI-----PNFLVDMISPAQITMLPHLIVDVRLTLEIM 155  
Db 855 YWVLSHGERESKUGSATIVNITILKNDDPHGIIIEFVSDGLVIMNESKGA----- 907  
QY 156 TPLDVRNVYHFNNQOPRMLVAMLYTPLRNSGDDVFTVSCRVLTRPTPDP----- 210  
Db 908 -----IYSAVDVVRNRCNFGDV-SVSWVV-----SPDFTQDVF 940  
QY 211 -----EF-----IYLVPPSVESKTPFTLPIL-----TISELTNSRPP 243  
Db 941 PVQGTWVFGDQEFKSNITVSLPDEIPEEMBEFTVILLNGTGAKVGNRTATLIRRID 1000  
QY 244 IPIEOLYTPNETVNVQONGRCTLGDELQGTQTLSSAV-----CFLOGRVADNGDNWDQ 300  
Db 1001 DPI-----YFA--EPRVVRVQ-----EGETANFTVLRNGSDVTCMVQ----- 1036  
QY 301 NLLQLTPNGASYDPTDEVAPLPGTQDFSGMLYGV-LTQDNVNVSTGE-----AKNAKGIYI 356  
Db 1037 ---YATKQKATAREDRIPVEK-----ETLIFEVGSQQSISIFVNEGDGIPETDEPYI 1089  
QY 357 STTSKFTPKTIGSLHSITEHVH--PNOQSRFTPVGVAVDENTPFQOWVLPH--YAGSL 412  
Db 1090 ILLNSTGDTVVYQVATVITEANDDPNGIFSLPIEDKAVEEGKTNFWRILHRHRYFGSV 1149  
QY 413 ALNTNLAPAVATPPGEQLLPRSRVPCVQOGQODAFIDCLLPQEWNVHFYQEAAPSOA 472  
Db 1150 SVSWQLFQNDALQPGQE----FYETSGTNFMDGSEA-----KPIILHAFPPDKIPBFN 1199  
QY 473 DVALRYVN-----PDTGRITFEAKLHRSFGFTVTSHTGAYPLVWPNGHFRFD 520  
Db 1200 EYFELKLVNISGSGPQGGQLAETNLQ-----VTV-----WVPNDPDPFGVIFLD 1244

## RESULT 15

US-10-389-566-1196  
; Sequence 1196, Application US/10389566  
; Publication No. US20040025202A1  
; GENERAL INFORMATION:  
; APPLICANT: Monsanto Technology, LLC  
; APPLICANT: Laurie, Cathy C  
; TITLE OF INVENTION: Nucleic Acid Molecules Associated with Oil in Plants  
; FILE REFERENCE: 38-77(52900)D  
; CURRENT APPLICATION NUMBER: US/10/389,566  
; CURRENT FILING DATE: 2003-03-31  
; PRIOR APPLICATION NUMBER: US 60/365,301  
; PRIOR FILING DATE: 2002-03-15  
; PRIOR APPLICATION NUMBER: US 60/391,786  
; PRIOR FILING DATE: 2002-06-25  
; PRIOR APPLICATION NUMBER: US 60/392,018  
; PRIOR FILING DATE: 2002-06-26  
; NUMBER OF SEQ ID NOS: 2459  
; SOFTWARE: Patent in version 3.2  
; SEQ ID NO 1196  
; LENGTH: 1109  
; TYPE: PRT  
; ORGANISM: Ipomoea nil  
US-10-389-566-1196

Query Match 3.8%; Score 109.5; DB 16; Length 1109;  
Best Local Similarity 19.7%; Pred. No. 1.4;  
Matches 130; Conservative 84; Mismatches 184; Indels 261; Gaps 37;  
QY 9 APSNDGAASL-----VPEGINETMPLEPVAGNSIAAPVAGQNTIIDPWI-----RTN 55  
Db 23 ALNSDGAALLSLRHTWTSIPSDITQSW-----NASDSTPCS-----WLGVECDRRQ 68  
QY 56 FVQAPNGEFTVSPRNSPCEILLNLELGPDLNPNYLAHLRMY-----NGVAGGVEVQ----- 106  
Db 69 FVDILN-----LSSYG---ISGEFGPEIS-HLKLKKVLSVNGFGSGPSQLGNCSL 117  
QY 107 ---VLLAGNATAGKILFAAIPPN-----FLVDMIS--PAQITMLPHL----- 144  
Db 118 LEHIDLSSNFTGN-----IPDTLGAQLNLRNLSPFNLSLIGPPFESLLSIPHLETVVF 171  
QY 145 -----IVDVRTL-----EPIMTLPD-----VRNVFVFNH-----QRP 174  
Db 172 TGNGLNGSIPNIGMNSLTTLLDDNQFSGVPVSSLGNITTLQSLYLNDNNLVGLTLPVT 231  
QY 175 RMRIVAMLYTPLRNSG--SGDDVFTVSCRVL-TRPTDPFEFTYLVPPSVESKT--KPPTL 229  
Db 232 LNNLENVYLDVRNNSLVGAIPDFVSCQKQDITISLNNQFTGGLPGLGNCISUREGA 291  
QY 230 PILTISELTNSRFPPI-----IEQLYTAPNE-TNVVQONGRC-----TLD 269  
Db 292 FSCALSG-----PIPCFGQLTKLDTLYLAGNHFSGRIPPELGCKSMIDLQLOQNOLE 345  
QY 270 GELOTTTOLLS-----SVCFLQ----- 287  
Db 346 GEIPGELGMLSQLQYLHLYTNLNGEVPISLWIKQSLQSLQLYQNLSGELPVDMTLQK 405  
QY 288 -----GRTVADNGDNWDQNLILQLTV-----PN-----CASVDP 315  
Db 406 LVSLALYENHTGVIPQDLGANSSELEVLDLTRNFTGHIIPNLCSQKLLKLLGYNY-L 464  
QY 316 TDEVPAPL-----GTQDF--SGMLYGVLTQDN-----VNVSTGEARN 350  
Db 465 EGSVPDLGGCCTLERLILEBNNLRGGLPDPVEKQNLFFDLSGNNFTGPIPPSLGNLKN 524  
QY 351 AKGIYISTT--SGKFTPKIGSIG-----LHSTEHVHPNQOQSRFTPVGVAVDENTPFQ 401  
Db 525 VTAYLSSNLQSGSIPPELGLSVLKEHLNHLNHLKGLPSELSELSNCHKLS-ELDASHNLL 583  
QY 402 QWVLPHYAGS-----LALNTNLAPAVAPT--EPGEQLL-----FFRSRVPCVOGLQ 445  
Db 584 NGSIPSTLGSITELTKUSLGNSFSGGIPTSLFSQSNKULNLQLGNNLAGDIPPVGALQ 642

us-09-926-799-9.rapb

Wed Jun 2 09:13:34 2004

Search completed: June 1, 2004, 14:04:44  
Job time : 36.648 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.54765 Seconds  
(without alignments)  
4310.252 Million cell updates/sec

Title: US-09-926-799-9  
Perfect score: 2863  
Sequence: 1 MKNASNDAPSDNGAASLVP.....VNQFSLAPMTGNGRRRIQ 542

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1830.5	63.9	539	1 COAT LORDV	P54635 lordedale v
2	1159.5	40.5	546	1 COAT SOUV3	Q04542 southampton
3	340	11.9	2344	1 POLN RHV	P27410 rabbit hemo
4	281.5	9.8	2208	1 POLN WANCV	Q69014 manchester
5	279	9.7	668	1 COAT FCVC6	P27404 feline cali
6	276.5	9.7	671	1 COAT FCVF9	P27406 feline cali
7	271.5	9.5	668	1 COAT FCVF4	P27405 feline cali
8	269.5	9.4	702	1 COAT SMSV1	P36284 san miguel
9	240	8.4	703	1 COAT SMSV4	P36285 san miguel
10	145.5	5.1	2333	1 POLG FMDV1	P03306 f genome po
11	130	4.5	2332	1 POLG FMDVA	P03308 f genome po
12	129	4.5	2193	1 POLG CX16T	Q9qf31 c genome po
13	123.5	4.3	746	1 RHTA RHIME	Q923q5 rhizobium m
14	118	4.1	2194	1 POLG HE701	P23537 h genome po
15	116.5	4.1	2193	1 POLG CX16G	Q65900 c genome po
16	116	4.1	545	1 GVPD HALME	Q02229 halobacteri
17	114.5	4.0	725	1 SP3 MOUSE	O70494 mus musculus
18	114	4.0	2336	1 POLG FMDVZ	P49303 f genome po
19	111.5	3.9	3317	1 CADN RAT	P58365 ratus norv
20	111	3.9	1034	1 ADD DROME	P54362 drosophila
21	110.5	3.9	5376	1 ZAN MOUSE	O88799 mus musculus
22	110	3.8	1196	1 XPG XENLA	P14629 xenopus lae
23	110	3.8	2290	1 POLG EMCV	P03304 c genome po
24	109.5	3.8	808	1 POLG HPAGV	Q02381 hepatitis a
25	109.5	3.8	852	1 POLG HPAPV	P06442 hepatitis a
26	109.5	3.8	2193	1 POLG HE71M	Q66479 h genome po
27	108.5	3.8	901	1 POLG ENMG3	P32540 mengo encep
28	108.5	3.8	2226	1 POLG HPAP2	P26580 hepatitis a
29	108.5	3.8	2226	1 POLG HPAP4	P26581 hepatitis a
30	108.5	3.8	2227	1 POLG HPAPV	P08617 hepatitis a
31	108.5	3.8	2227	1 POLG HPAPV	P06441 hepatitis a
32	108.5	3.8	2227	1 POLG HPAPM	P13901 hepatitis a
33	108.5	3.8	2230	1 POLG HPAPV	P14553 simian hepa

```

34 108 3.8 2193 1 POLG HE71B Q66478 h genome po
35 108 3.8 2332 1 POLG FMDVO P03305 f genome po
36 107.5 3.8 2175 1 POLG BOVEV P12915 b genome po
37 107 3.7 2206 1 POLG POL3L P03302 poliovirus
38 106.5 3.7 781 1 SP3 HUMAN Q02447 homo sapien
39 106.5 3.7 2226 1 POLG HPAPV P26582 hepatitis a
40 105 3.7 596 1 VP3A_HUMAN Q96ax1 homo sapien
41 104 3.6 602 1 SYD_NSIIMB Q9kous neisseria m
42 104 3.6 2292 1 POLG EMCVB P17593 e genome po
43 103.5 3.6 839 1 POLG HPAPV P31788 simian hepa
44 103 3.6 834 1 POLG ENMGO P12296 mengo encep
45 102.5 3.6 507 1 VL2_HPVI1 P03305 human papil

```

RESULT 1

COAT LORDV STANDARD; PRT; 539 AA.

AC P54635;

DT 01-OCT-1996 (Rel. 34, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-NOV-1997 (Rel. 35, Last annotation update)

DE Coat protein (Capsid protein).

OS Lordedale virus (Human enteric calicivirus).

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.

OX NCBI\_TaxID=82658;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=96005060; PubMed=7561776;

RA Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;

RT "Human enteric Caliciviridae: the complete genome sequence and expression of virus-like particles from a genetic group II small round structured virus".

RL Gen. Virol. 76:2349-2355 (1995).

CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@sib-sib.ch).

CC EMBL; X86557; CAA60255.1; .

DR InterPro; IPR004005; Calici\_coat.

DR InterPro; IPR008975; Viral\_cap\_coat.

DR Pfam; PF00915; Calici\_coat; 1.

KW Coat protein; Glycoprotein.

FT CARBOHYD 172 172 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 307 307 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 406 406 N-LINKED (GLCNAC. .) (POTENTIAL).

FT CARBOHYD 531 531 N-LINKED (GLCNAC. .) (POTENTIAL).

SQ SEQUENCE 539 AA; 58775 MW; 5E5C63E7F2C5FD21 CRC64;

Query Match 63.9%; Score 1830.5; DB 1; Length 539;

Best Local Similarity 64.0%; Pred. No. 4e-129;

Matches 349; Conservative 73; Mismatches 110; Indels 13; Gaps 6;

```

QY 1 MKNASNDAPSDNGAASLVP...VNQFSLAPMTGNGRRRIQ 542
Db 1 MKNASNDAPSDNGAASLVP...VNQFSLAPMTGNGRRRIQ 542
QY 61 NGEFTVSPRNSGPEILLNLELPGDLPNLYLAHLSRMVYAGVGVQVLLAGNAFTAGKTL 120
Db 61 GGEFTVSPRNPAGETLWSAPLPGDLPNLYLAHLSRMVYAGVGVQVLLAGNAFTAGKVI 120
QY 121 FAATPPNPLVDMSPAQITMLPHLIVDRTLEPINTPLPDVNVFVHNNQPPRMRLVA 180
Db 121 FAAPPNPFPTEGLSPQVTFPHIIVDRQLFVLIPLPDVNRNPFHYNQANDSTLKLA 180

```

Matches 258; Conservative 84; Mismatches 176; Indels 51; Gaps 17;

QY 1 MLYTPIRNSGDDVFTVSCVRLTRPTDFFIYLVPPSVESKTKPFTLPILITSELTNS 240  
 DB 1 MLYTPIRANAGDDVFTVSCVRLTRPTDFFIYLVPPSVESKTKPFTVPVLTVESNS 240  
 QY 241 RPIPIEQLYTAAPNETVNVQCGRCTLGDELQGTLLSSAVCFLOQRTVADNGDNDQ 300  
 DB 241 RPIPIEQLYTGSSAFVVPQNGRCTLGDELQGTLLSSAVCFLOQRTVADNGDNDQ 298  
 QY 301 NLLQLYTPNGASYDTPDEVPALGTQDSGMLYGVLTQDNVNVSTGEAKNAKGIYISTTS 360  
 DB 299 YTMNLASQWNSYDTPDEVPALGTQDSGMLYGVLTQDNVNVSTGEAKNAKGIYISTTS 355  
 QY 361 GKFTPKIGISGLHSITEH-VHENQOSRFTFPVGAVD-----ENTPQQWVLFHYAGSLAIN 415  
 DB 356 VHTFPKLSGVQFTTDNNDFQAGQNTKFTFPVGIQDGHQNEP-QQWSLPNYSGRTHN 414  
 QY 416 TNLAPAVATPFGQELLFRSRVPCVQGLQGDADFIDCLLPQEWVNFYQEAAPQSDVA 475  
 DB 415 VHLAPAVATPFGQELLFRSRVPCVQGLQGDADFIDCLLPQEWVNFYQEAAPQSDVA 472  
 QY 476 LTRYNPDGRTLFEAKLHRSGLTSHYTGAYPLVVPNGHFRFDSWVNFYSLAPMGTVG 535  
 DB 473 LLRFVNPDTGRVLFCKLHRSGLTSHYTGAYPLVVPNGHFRFDSWVNFYSLAPMGTVG 532  
 QY 536 NGRRR 540  
 DB 533 TGRRR 537

RESULT 2  
 COAT\_SOUV3 STANDARD; PRT; 546 AA.  
 ID COAT\_SOUV3  
 AC Q04542:  
 DT 01-OCT-1994 (Rel. 30, Created)  
 DT 01-OCT-1994 (Rel. 30, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS Southampton virus (serotype 3).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OC NCBI\_TaxID=37129;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=93142023; PubMed=8380940;  
 RA Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;  
 RT "Sequence and genome organization of a human small round-structured  
 RT (Norwalk-like) virus."  
 RL Science 259:516-519(1993).  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; L07418; AAA92984.1; -.  
 DR PIR; B37491; B37491.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;  
 Query Match 40.5%; Score 1159.5; DB 1; Length 546;  
 Best Local Similarity 45.3%; Pred. No. 5.2e-79;

MATCHES 258; Conservative 84; Mismatches 176; Indels 51; Gaps 17;

QY 1 MLYTPIRNSGDDVFTVSCVRLTRPTDFFIYLVPPSVESKTKPFTLPILITSELTNS 240  
 DB 1 MLYTPIRANAGDDVFTVSCVRLTRPTDFFIYLVPPSVESKTKPFTVPVLTVESNS 240  
 QY 241 RPIPIEQLYTAAPNETVNVQCGRCTLGDELQGTLLSSAVCFLOQRTVADNGDNDQ 300  
 DB 241 RPIPIEQLYTGSSAFVVPQNGRCTLGDELQGTLLSSAVCFLOQRTVADNGDNDQ 298  
 QY 301 NLLQLYTPNGASYDTPDEVPALGTQDSGMLYGVLTQDNVNVSTGEAKNAKGIYISTTS 360  
 DB 299 YTMNLASQWNSYDTPDEVPALGTQDSGMLYGVLTQDNVNVSTGEAKNAKGIYISTTS 355  
 QY 361 GKFTPKIGISGLHSITEH-VHENQOSRFTFPVGAVD-----ENTPQQWVLFHYAGSLAIN 415  
 DB 356 VHTFPKLSGVQFTTDNNDFQAGQNTKFTFPVGIQDGHQNEP-QQWSLPNYSGRTHN 414  
 QY 416 TNLAPAVATPFGQELLFRSRVPCVQGLQGDADFIDCLLPQEWVNFYQEAAPQSDVA 475  
 DB 415 VHLAPAVATPFGQELLFRSRVPCVQGLQGDADFIDCLLPQEWVNFYQEAAPQSDVA 472  
 QY 476 LTRYNPDGRTLFEAKLHRSGLTSHYTGAYPLVVPNGHFRFDSWVNFYSLAPMGTVG 535  
 DB 473 LLRFVNPDTGRVLFCKLHRSGLTSHYTGAYPLVVPNGHFRFDSWVNFYSLAPMGTVG 532  
 QY 536 NGRRR 540  
 DB 533 TGRRR 537

RESULT 3  
 POLN RHDV STANDARD; PRT; 2344 AA.  
 ID POLN RHDV  
 AC P27410;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Non-structural polyprotein [Contains: RNA-directed RNA polymerase  
 DE (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like  
 DE protein); Coat protein].  
 OS Rabbit hemorrhagic disease virus (RHDV).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Lagovirus.  
 OC NCBI\_TaxID=11976;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91361557; PubMed=1840711;  
 RA Meyers G., Wirblich C., Thiel H.-J.;  
 RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide  
 RT sequencing of a calicivirus genome."  
 RL Virology 184:684-676(1991).  
 CC -!- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE  
 CC FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC [RNA] (N).  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

DR EMBL; M67473; AAA47285.1; -  
DR PIR; A41039; RWRWH.  
DR PDB; 1KHV; 16-JAN-02.  
DR PDB; 1KHV; 16-JAN-02.  
DR MEROPS; C24.001; -  
DR InterPro; IPR003593; AAA ATPase.  
DR InterPro; IPR004005; Calici coat.  
DR InterPro; IPR004004; Calici\_pol\_hel.  
DR InterPro; IPR000317; Peptidase\_C24.  
DR InterPro; IPR000605; RNA helicase.  
DR InterPro; IPR007095; RNA pol DS PS.  
DR InterPro; IPR001205; RNA pol p3D.  
DR InterPro; IPR007094; RNA\_pol\_Psvir.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
DR Pfam; PF03510; Peptidase\_C24; 1.  
DR Pfam; PF00680; RNA dep RNA\_pol; 1.  
DR Pfam; PF00910; RNA helicase; 1.  
DR PRINTS; PR00916; 2CENDOPTASE.  
DR PRINTS; PR00918; CALICIVIRUSNS.  
DR SMART; SM00382; AAA; 1.  
KW Polyprotein; Transferrase; RNA-directed RNA polymerase; Hydrolase;  
KW Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.  
FT CHAIN 7508 7707  
FT CHAIN 1109 1251  
FT CHAIN 71497 71625  
FT CHAIN ? 2344  
FT CHAIN ? 2344  
FT NP\_BIND 522 539  
FT ACT\_SITE 1212 1212  
FT ACT\_SITE 1227 1227  
FT ACT\_SITE 1227 1227  
SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 11.9%; Score 340; DB 1; Length 2344;

Best Local Similarity 27.5%; Pred. No. 5.9e-17;

Matches 145; Conservative 72; Mismatches 213; Indels 98; Gaps 25;

QY 8 AAPSNDGAASL-----VPEGINETMLEPVA-----GASIA-APVAGQTNIID--- 49  
DB 1772 AAQGEAAAGTATTASVPGTTDGMPPGVVATTSVITAENSSAIATAGIGPPQVDOQE 1831  
QY 50 PHIRTNFOAPNGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVEVQVLL 109  
DB 1832 TW-RNFYY--NDVFTWSVADAPGSLITYVQHSPPQNPPTAVLSQWYAGWAGMGRFIV 1888  
QY 110 AGNAFTAGKILFAAIPPNFLVDIMISPAQITMLPHLIVDVRTLEIMTLPDVRNVFYHFN 169  
DB 1889 AGSGVFGGRLVRAVPPG--IEIGPGLVRQFPVHVVIDARSLPEVITITMPDLRPNMYHT 1946  
QY 170 NOPQRMRLVAMLYTPLSNGSGDDVFTVSCVLRTPDPDFEYLVPSPVESKTPFTL 229  
DB 1947 GDFGLVPLTLVSVYNNL-INPFGSGTSATQVTVTRPSEDFVMIRAPS--SKTVDSIS 2003  
QY 230 P--ILITISLT---NSRPEIPDLYTAPNETNVVQCONGRCTLDGELQGTTLSSAV 283  
DB 2004 PAGLUTTPVLTGVGNDNRWNGQIVGLQVPVGGFST--C-NRHNHNGSTYGNSSPRFGDI 2060  
QY 284 CFLQRTVADNGDNDQNLQLTYPNGASY--DPTDEVPAPLGTQD-----PSG----- 330  
DB 2061 DHRRG-SASYSGSN-ATNVLPQWYANAGSAIDNPISQV-APDGFDPMSVPVNGPGIPAA 2117  
QY 331 --MLYGVLTQDN---VNVSTGEAKN-----ANGYIYSTT--- 359  
DB 2118 GWVGFCAIWNNSGAPNVTTVQAYELGATGAPGNLQPTTNTSGATVAKSIYAVVTGTA 2177  
QY 360 ---SKGFTPKIGSLGHSITEVHFNQOQSRF-----TPGVAVDENTPQWVLPYHAGS 411  
DB 2178 QNPAGLFWNASGIISTPNASAITYTPQDRIVTTTGTGTPAAAPVGNKTPIM-----FAS 2230

QY 412 LALNTNLAPAVAPTTPGGQLLFFRRVPCVQGLQGQDAPIDCLLPQEW 459  
DB 2231 VVRRTGDNATAGSANGTYGTGSQLPVTIGL-SLNNYSSALMEQGF 2277

#### RESULT 4

POLN\_MANCV STANDARD; PRT; 2208 AA.  
ID POLN\_MANCV AC Q69014;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polyprotein [Contains: RNA-directed RNA polymerase  
DE (EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like  
DE protein); Coat protein].  
OS Manchestera virus (Human enteric calicivirus).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Sapovirus.  
OX NCBI\_TaxID=82659;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=95390791; PubMed=7661689;  
RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;  
RT "Human enteric caliciviruses have a unique genome structure and are  
RT distinct from the Norwalk-like viruses."  
RL Arch. Virol. 140:1345-1356(1995).  
CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).  
CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR  
CC THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY  
CC SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC {RNA}(N).  
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch))

CC EMBL; X86560; CAA60262.1; -  
DR MEROPS; C24.UPW; -  
DR InterPro; IPR004005; Calici coat.  
DR InterPro; IPR004004; Calici\_pol\_hel.  
DR InterPro; IPR000317; Peptidase\_C24.  
DR InterPro; IPR000605; RNA helicase.  
DR InterPro; IPR007095; RNA pol DS PS.  
DR InterPro; IPR001205; RNA pol p3D.  
DR InterPro; IPR007094; RNA\_pol\_Psvir.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
DR Pfam; PF03510; Peptidase\_C24; 1.  
DR Pfam; PF00680; RNA dep RNA\_pol; 1.  
DR Pfam; PF00910; RNA helicase; 1.  
DR PRINTS; PR00916; 2CENDOPTASE.  
DR PRINTS; PR00918; CALICIVIRUSNS.  
KW Polyprotein; Transferrase; RNA-directed RNA polymerase; Hydrolase;  
KW Thiol protease; Helicase; ATP-binding; Coat protein.  
FT CHAIN ?  
FT CHAIN ?  
FT CHAIN ?  
FT CHAIN ?  
FT NP\_BIND 408 415  
FT ACT\_SITE 1097 1097  
FT ACT\_SITE 1112 1112  
FT ACT\_SITE 1112 1112  
SQ SEQUENCE 2208 AA; 242736 MW; 3E299DSBA20E45DC CRC64;

Query Match 9.8%; Score 281.5; DB 1; Length 2208;

Best Local Similarity 21.7%; Pred. No. 1.3e-12;

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

Matches 128; Conservative 92; Mismatches 228; Indels 143; Gaps 22;  
QY 20 PEGINETMPLFVAGASIAAPVAGQTNIDPWIRTNFVQAPNGEFTVSPRNSGEILLNL 79  
Db 1687 PNGAQRLELAATGAI-----QSNVPEA-IRNCF--AVFRTFAWDRMPTGTFGLGSI 1736  
QY 80 ELGPDNLNPLYLAHLSRMVNGYAGGVQVLLAGNATAGKILFAAIPNPNFLVDMISPAQIT 139  
Db 1737 SLHPNINPNTSLSGMWAGSGSFVRLSISGSGVFAGRITIASVIPPQ--VDPSSIRDPG 1794  
QY 140 MLPHLIVDVRTLEPTMTPLDVNRYFHNQOPQPRMELVAMLYPLPLNSGSGDDVTVS 199  
Db 1795 VLPHAFVDAIRIPEVSFMDPVRAVDYHRMDGAEPCTSLGFWVYQPLNPFSTTAVSTCW 1854  
QY 200 CRVLTRPTDFEYIYVLP--SVESKTKP-----FTLPIILTISE--L 237  
Db 1855 VSVETKPGGDFDFCLLRPPGQMGVSGEGLLPRLGYSRGNRVGGLVGMILVAEHKQ 1914  
QY 238 TNSRF-----PIEQLYTAPNETN-----VVQCN-----G 264  
Db 1915 VNRHNSNVTFCWSTAPVNPMAAEIVTQAHSTSRHAWLSIGAQNKGPLFFGIPNHPD 1974  
QY 265 RC-----TLDBELQ--TTQLSSAVCFLOQRTVADGNDQNLQLQTYENGASYDPT 316  
Db 1975 SCATVVGAMDTSLGRPSTGVCPSAIF-----QNNGDVYENDTTPSVMF--ATYDPL 2025  
QY 317 DEVPAPIGTQDFSGMLYGLVTQDNVNVSTGEAKNAKGIYISTTSKGTPKIGSLGHSIT 376  
Db 2026 TSGTGVALTNSINPASILVRIENDEFTSGFANDKNVVQMSWEMVT-----GTNQIR 2079  
QY 377 EHVHPNQSRFTPVGVAVDENTFPQWVLPHYAGSLALNTNLAPAVAPTFPGEQLLFFRS 436  
Db 2080 GOVTPMSGTNTFTSTCANTLVLMQERML-----SYDGHQALYSS 2120  
QY 437 RVPCVQGLQGDQAFIDCLLQPEWVNFYQEAAPSQADVALIRVYNPDRTGRTLFEAKLHRS 496  
Db 2121 QL-----ERTAEYFQNDIVNIPNSMAVF-NVETNSASFQIGIRPD 2160  
QY 497 GFITVSHTGAYPLVVPNGHFRFDSWVNFYSLAPMG-----TGN-GRRR 540  
Db 2161 GMYVTG--GSGVNVPLEPETRF-----QYVGLPLSALSGFSGNMGRRAR 2204  
RESULT 5  
COAT FCVC6 STANDARD; PRT; 668 AA.  
AC COAT FCVC6  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Coat protein (Capsid protein).  
OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Vesivirus.  
OX NCBI\_TaxID=11979;  
RN [1]  
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RX MEDLINE=91374597; PubMed=1716692;  
RA Neill J.D., Reardon I.M., Heinrikson R.L.;  
RT "Nucleotide sequence and expression of the capsid protein gene of  
feline calicivirus."  
RL J. Virol. 65:5440-5447(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Neill J.D.;  
RL Submitted (Sep-1994) to the EMBL/GenBank/DBJ databases.  
CC -1- PTM: The N-terminus is blocked.  
CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
CC -1- This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way

Query Match 9.7%; Score 279; DB 1; Length 668;  
Best Local Similarity 23.2%; Pred. No. 3.8e-13;  
Matches 137; Conservative 81; Mismatches 210; Indels 162; Gaps 28;  
QY 11 SNGGAASLVPEGINETMP-----LEPVAGASIAAPVAGQTNIDPW-IRTNFVQAPNGEFT 65  
Db 125 ADDGSITTEPQG---TMVGGVIAEPNAQMSAADWATGKSVDSWEAFSPHTSVNW--- 178  
QY 66 VSPRNSPGEILLNLEGLPDNLNPLYLAHLSRMVNGYAGGVQVLLAGNATAGKILFAAIP 125  
Db 179 -STSETQKILFKQSLGGLNLPYLTLAKLYAVWSGVSDVRFSGSGVFGKJAAIVVP 237  
QY 126 PNFVDMISPAQITMLPHLIVDVRTLEPTMTPLDVNRYFHNQOPQPRMELVAMLYTP 185  
Db 238 PG--IDPVQSTMSLQVPHVLFDAQVPEPVIFSDILRSLRSLYHLMSD--TTSIVIMVYND 294  
QY 186 L-----RNSGSGDDVFTVSCRVLTPTPTDFEYIYVLP-----SVESK----- 223  
Db 295 LINPYANDSNSGCV-----TVETKPGDPFKFHLKPPGSMTHGSIIPSDILPKSSSLW 349  
QY 224 -----TKPFTL-----PI-LTISELNSRFPPIEQ 248  
Db 350 IGNRFWSDITDFVIRPFPVQANRHFDFNQTAGMSTFRPRPITITISVKESAKLGIGVAT 409  
QY 249 LYTAPN-----ETNVVQCNGRCCTLDGELQGTITQLLSAVCFLOQRTVADGNDQNL 302  
Db 410 DYIVEGIPDGMPDT-----TIPGELVPV-----GDYAITNGTND--- 444  
QY 303 LQLYTPNGASYDPTDEVPAPLGTQDFSGM-LYGVLTQ-----DNVNVSTGEAKNAKG 353  
Db 445 -----ITTAQYDAATEI---RNNTNFRGMYTCGSLQRAWGDKKISNTAFITGIVDQAKL 497  
QY 354 IYISTTSKGTPKIGSLGHSITTEHVHNQOS-----RFTPVG-----VAVDENTFPQWVLP 406  
Db 498 IPSNTID---QTKIAVFOOTIANKHVQTSDDTLALLGYTGIGEEAIGADRVRVIRSLVP 554  
QY 407 HYAGSLALNTNLAPAVAPTFPGEQLLFFRSRVPCVQGLQGDQAFIDCLL-----PQEWVNF 463  
Db 555 ERGAR-----GGNHPIFKNSIKLGYVIRSDVFNQSILHTSRQLSLNHY 599  
QY 464 YQEAAPSQADVALIRVYNPDRTGRTLFEAKLHRS--FTTVSHTG--AYPL 509  
Db 600 LL-----SPDSFAVYRII---DSNGSWFIDIGDNDGFSFGVSGSIGKLEFPL 643  
RESULT 6  
COAT FCVC9 STANDARD; PRT; 671 AA.  
ID COAT FCVC9  
AC P27406;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE Coat protein (Capsid protein).  
GN CPP76.

OS Feline calicivirus (strain F9) (FCV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OX NCBI\_TaxID=11981;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92410623; PubMed=1529544;  
RA Carter M.J., Milcon I.D., Meanger J., Bennett M., Gaskell R.M.,  
RT Turner P.C.;  
RL "The complete nucleotide sequence of a feline calicivirus";  
RN Virology 190:443-448(1992).  
[2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92339673; PubMed=1633955;  
RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;  
RT "Cloning and sequence determination of the feline calicivirus strain  
F9";  
RL Biochem. Soc. Trans. 20:265-268(1992).  
[3]  
CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL; M86379; AAA79327.1; -;  
DR EMBL; Z11536; CA77636.1; -;  
DR PIR; B43382; VCMWF9.  
DR InterPro; IPR004005; Calici\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
KW Coat protein; Glycoprotein.  
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 671 AA; 73441 MW; 338E86D8370D5E5 CRC64;  
Query Match 9.7%; Score 276.5; DB 1; Length 671;  
Best Local Similarity 25.1%; Pred. No. 5.8e-13;  
Matches 131; Conservative 68; Mismatches 171; Indels 151; Gaps 28;  
QY 11 SNDGAASLVPGINETMP-----LEPVAGASTAAPVAGQTNIIDPW-IRTNFVQAPNGEFT 65  
DB 125 ADDGSITAEQ---TWVGGVIAEPSAQMSMAADMATKSVDSWEAFSPHTSVNW---178  
QY 66 VSPRNSPGEILLNLGLPDLNLPYLAHLSRMVYAGGVGVQVLLAGNAFTAGKILFAAIP 125  
DB 179 -STSETQKILFKQSLGLPPLNYPYLAHLSRMVYAGGVGVQVLLAGNAFTAGKILFAAIP 237  
QY 126 PNFLVDMISPAQITMLPHLIVDRTLEPIMTLPDVRNVFHNQPPMRRLVAMLYTP 185  
DB 238 PG--VDPVQSTSMQLQYPHVLEDAQVEPVIIFCLPLRLSTLYHLSMDTD--TTSLVIMVYND 294  
QY 186 LRSNGSGDDVFTVSC--RVLRTPDPDEFIYVLP-----SVES-----KT-----224  
DB 295 L-INPYANDANSSGCVITVETKPGDPFKHLLKPPGSMVTGHSIPSLDIPKSPSLWIGNR 353  
QY 225 -----KPFLL-----PI-LTISELTNSRRFPIEQLYTA 252  
DB 354 YWSDITDFVIRPFVQANRHFDFNQETAGMTSPRPRISVITEQNGAKLIGVATDIV 413  
QY 253 PN-----ETNVVQCONGRCTLDGEL-----QGTQLLSSAVCFQGRVAD---N 294  
DB 414 FGIPDGPEDT-----TIPGELIPAGDYAITNGTNDITTAT---GYDADIKN 459  
QY 295 GDN-----W-DQMLLQITYNGASYD--PTDEVAPLGTQDFSGMLVGLVLTQD 339  
DB 460 NTNFRMGYICGSLQRAWGDKKISNTAFITTTATLDGDNNNKINPCNTIDQSKI---VFPQD 516

QY 340 N-----VNVG-----TGEAKNKG-----IYISTSGKTPKIGTGLHSITHV 379  
DB 517 NHVGKKAQTSDDTLALLGYTGIGGQAIGSDRDRVVRIST-----LPETGARGGN-----565  
QY 380 HPNQSRRFTPGVAVDENTPFQWVLPYHAGSLAINTNLAP 420  
DB 566 HPIFYKNSIKUGYVIRSIDVFNSQL-HTSQSLNHYLLP 605

## RESULT 7

COAT\_FCVF4  
ID COAT\_FCVF4 STANDARD; PRT; 668 AA.  
AC P27405;  
DT 01-AUG-1992 (Rel. 23, Created)  
DT 01-AUG-1992 (Rel. 23, Last sequence update)  
DT 01-JUN-1994 (Rel. 29, Last annotation update)  
DE Coat protein (Capsid protein).  
OS Feline calicivirus (strain Japanese F4) (FCV).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OX NCBI\_TaxID=11980;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91306470; PubMed=1853578;  
RA Tohya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,  
RA Miyamura K., Yamazaki S., Mikami T.;  
RL "Sequence analysis of the 3'-end of feline calicivirus genome";  
RL Virology 183:810-814(1991).  
CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
DR EMBL; D90357; BAA14371.1; -;  
DR PIR; B40481; VCMWF9.  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
KW Coat protein; Glycoprotein.  
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 668 AA; 73589 MW; 85BBD8C85804E503 CRC64;

Query Match 9.5%; Score 271.5; DB 1; Length 668;  
Best Local Similarity 22.5%; Pred. No. 1.4e-12;  
Matches 130; Conservative 77; Mismatches 193; Indels 179; Gaps 25;

QY 30 EPVAGASTAAPVAGQTNIIDPW-IRTNFVQAPNGEFTVSPRNSPGEILLNLGLPDLNYPY 88  
DB 145 EPSAQMSMAADMATKSVDSWEAFSPHTSVNW-----STSETQKILFKQSLGLPPLNYPY 200  
QY 89 LAHLSRMVYAGGVGVQVLLAGNAFTAGKILFAAIPENFLVDMISPAQITMLPHLIVDV 148  
DB 201 LEHLSKLYVAVWSGIEVRSISGSGVFGKLAIVVPPG--VDPVQSTSMQLQYPHVLFDA 258  
QY 149 RTLEPIMTLPDVRNVFHNQPPMRRLVAMLYTP-----RNSGSGDDVFTVSCRV 202  
DB 259 RQVEPFIITPDLRLSTLYHLSMDTD--TTSLVIMVYNDLINPYANDNSNGCIV-----TV 312  
QY 203 LTRTPDPDEFIYVLP-----SVES-----TKPFTL-----229  
DB 313 ETKPGDPFKHLLKPPGSMVTGHSIPSLDIPKSSSLWIGNRYWTDITDFVIRPFVQANR 372

151 AESDPCGADIIVTEEQGTWVQQQVPAQSAALTTAAASTGKT--VDCW--TTFFSYHTA 206  
 60 PNGEFTVSPRNSPGEILLNLELGPDLNPLAHLASMYNGVAGVEQVOLLAGNAFTAGKI 119  
 207 VNWSTT---BAQKILFSRALSPELNPYLRLHSSLYSTWSGGIDVRFVSGSGVFGKL 262  
 120 LFAALPPNPLFVDMISPAQITMLPHLIVDVRTLEIMPTPLDVRNVFVHFNQOQPRMLV 179  
 263 AALIVPPG--TEPVESPTMLQYPHVLFDAQTEVIFTIPDIRKTLX--SMDDTDTTTLV 319  
 180 AMLYTPL--RNSGSGDDVTVSCRVLTPTDPFEIYLVPP-----SVESKTP----- 226  
 320 INVYNELINPEYQSEPKSSCSITVETRESSDTFSLKPPGSLKKGSIPLDLPRNSRH 379  
 227 -----FTLPLTIS-----ELTNSRFPPIEOLYTAPNETNVVQCGRECT 267  
 380 WNGNRWSTIDGFVQPVFQSNRHFDFDSTTTGWSPT---YYTPIEVTLEKLDRG--- 432  
 268 LDGELGGTTQLLSSAVCFLOGRTVADNGDNWQNLQLTYPNGASVDPDDEVPAPLGTQD 327  
 433 --GOYFKVTDTEKSLVPLP-----DGMPTDTI-----PT-AMTASNGYD 470  
 328 FSCMLYGVLTQDNVNVSTGEAKNAKGIYISTTSGKFTPKI--GSIGLHSEITHVHNPQNSR 386  
 471 YTVAYRI-----TNGTHFKGFYI---MGNLTTKVKGSDNLGETQQ---TSRTL 514  
 387 FTPVGVAVDENT 398  
 515 FASVGNKYDQNT 526

RESULT 9  
 COAT\_SMSV1 STANDARD; PRT; 702 AA.  
 ID COAT\_SMSV1  
 AC P36284;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS San Miguel sea lion virus (serotype 1) (SMSV 1).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OX NCBI\_TaxID=36406;  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=92410750; PubMed=1529644;  
 RA Neill J.D.;  
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of  
 San Miguel sea lion virus: identification of conserved and non-  
 conserved amino acid sequences among calicivirus capsid proteins.";  
 RL Virus Res. 24:211-222(1992).  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; M87481; AAA16217.1; -.  
 DR F1R; A48562; A48562.  
 DR InterPro; IPR004005; Calici coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 481 481 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 493 493 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 545 545 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DEE3D7 CRC64;  
 Query Match 9.4%; Score 269.5; DB 1; Length 702;  
 Best Local Similarity 25.0%; Pred. No. 2.1e-12;  
 Matches 108; Conservative 58; Mismatches 169; Indels 97; Gaps 21;  
 8 AAFSNDGAASLVPEGINEWTFPEPVAGAS-----IAAPVAGQNIID-PWIRTFVQ---A 59

151 AESDPCGADIIVTEEQGTWVQQQVPAQSAALTTAAASTGKT--VDCW--TTFFSYHTA 206  
 60 PNGEFTVSPRNSPGEILLNLELGPDLNPLAHLASMYNGVAGVEQVOLLAGNAFTAGKI 119  
 207 VNWSTT---BAQKILFSRALSPELNPYLRLHSSLYSTWSGGIDVRFVSGSGVFGKL 262  
 120 LFAALPPNPLFVDMISPAQITMLPHLIVDVRTLEIMPTPLDVRNVFVHFNQOQPRMLV 179  
 263 AALIVPPG--TEPVESPTMLQYPHVLFDAQTEVIFTIPDIRKTLX--SMDDTDTTTLV 319  
 180 AMLYTPL--RNSGSGDDVTVSCRVLTPTDPFEIYLVPP-----SVESKTP----- 226  
 320 INVYNELINPEYQSEPKSSCSITVETRESSDTFSLKPPGSLKKGSIPLDLPRNSRH 379  
 227 -----FTLPLTIS-----ELTNSRFPPIEOLYTAPNETNVVQCGRECT 267  
 380 WNGNRWSTIDGFVQPVFQSNRHFDFDSTTTGWSPT---YYTPIEVTLEKLDRG--- 432  
 268 LDGELGGTTQLLSSAVCFLOGRTVADNGDNWQNLQLTYPNGASVDPDDEVPAPLGTQD 327  
 433 --GOYFKVTDTEKSLVPLP-----DGMPTDTI-----PT-AMTASNGYD 470  
 328 FSCMLYGVLTQDNVNVSTGEAKNAKGIYISTTSGKFTPKI--GSIGLHSEITHVHNPQNSR 386  
 471 YTVAYRI-----TNGTHFKGFYI---MGNLTTKVKGSDNLGETQQ---TSRTL 514  
 387 FTPVGVAVDENT 398  
 515 FASVGNKYDQNT 526

RESULT 9  
 COAT\_SMSV4 STANDARD; PRT; 703 AA.  
 ID COAT\_SMSV4  
 AC P36285;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS San Miguel sea lion virus (serotype 4) (SMSV 4).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OX NCBI\_TaxID=36407;  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=92410750; PubMed=1529644;  
 RA Neill J.D.;  
 RT "Nucleotide sequence of the capsid protein gene of two serotypes of  
 San Miguel sea lion virus: identification of conserved and non-  
 conserved amino acid sequences among calicivirus capsid proteins.";  
 RL Virus Res. 24:211-222(1992).  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL; M87482; AAA16220.1; -.  
 DR PIR; C48562; C48562.  
 DR InterPro; IPR004005; Calici coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 89 89 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 208 208 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 329 329 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 463 463 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 482 482 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 703 AA; 77721 MW; C5DAD8223B261073 CRC64;

Query Match 8.4%; Score 240; DB 1; Length 703;  
 Best Local Similarity 29.3%; Pred. No. 3.3e-10;  
 Matches 63; Conservative 45; Mismatches 91; Indels 16; Gaps 8;

QY 8 AAPSNDGAASLPEG-----INSTMPLPVPAGASIAAPVAGQNIIDPWIRTNFVQAPNCE 63  
 DB 151 AESDQGSABIVTEQGTVVQQPAPAPATLATATATGK-SVQEQWM-TFF--SYHTS 206  
 QY 64 FTVSPRNSGEILLNLELDPNPNYLAHLSRMNGYAGVEVQVLLAGNAFTAGKILFAA 123  
 DB 207 INWSTVESQKILYSQALNPSINPLDHTAKLYSTWSGGIDVRFTVSGSGVFGKLAALL 266  
 QY 124 IPPNFLDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVYHFNNOQPQRMRLVAMLY 183  
 DB 267 VPPG--VEPIESVSLQYPHVLFDAQTPEVFTTIPDKIRKLFHSMDET-D-TKLVINPY 323  
 QY 184 TPLRNSGGDDVFTVSCRVLTPPTDPEFIVLVP 218  
 DB 324 ----ENGV-ENKTCSTIVETPSADFTALLKPP 353

RESULT 10  
 POLG\_FMDV1 STANDARD; PRT; 2333 AA.  
 AC P03306; Q84768; Q84750; Q84751; Q84752; Q84753; Q84754; Q84760;  
 AC Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;  
 AC Q84769; Q89824;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat  
 proteins VP1 TO VP4; Core protein P52; Genome-linked proteins VP61 TO  
 VP63; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed  
 DE RNA polymerase P56A (EC 2.7.7.48)]  
 OS Foot-and-mouth disease virus [strain A10-61] (Aphthovirus A) (FMDV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Aphthovirus.  
 OX NCBI\_TaxID=12112;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=84169547; PubMed=6324120;  
 RA Carroll A.R., Rowlands D.J., Clarke B.E.;  
 RT "The complete nucleotide sequence of the RNA coding for the primary  
 RT translation product of foot and mouth disease virus.";  
 RL Nucleic Acids Res. 12:2461-2472(1984).  
 RN [2]  
 RP SEQUENCE OF 115-1048 FROM N.A.  
 RX MEDLINE=82211814; PubMed=6282711;  
 RA Boothroyd J.C., Harris T.J.R., Rowlands D.J., Lowe P.A.;  
 RT "The nucleotide sequence of cDNA coding for the structural proteins  
 RT of foot-and-mouth disease virus.";  
 RL Gene 17:153-161(1982).  
 CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the  
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
 CC substituted for Gln, and Ser or Thr for Gly.  
 CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC (RNA) (N).  
 CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; V01130; CAA24361.1; -

DR EMBL; X00429; CAA25127.1; -.  
 DR MEROPS; C03.008; -.  
 DR MEROPS; C28.001; -.  
 DR InterPro; IPR004004; Calici\_pol\_hel.  
 DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
 DR InterPro; IPR004080; FMDVPicoat.  
 DR InterPro; IPR008739; Peptidase\_C28.  
 DR InterPro; IPR001876; RNv.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_P5vir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF03408; Peptidase\_C28; 1.  
 DR Pfam; PF00073; rhv; 3.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PR00918; CALICIVIRUSNS.  
 DR PRINTS; PR01542; FMDVPICOAT.  
 KW Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;  
 KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;  
 KW Myristate; Lipoprotein.  
 FT CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.  
 FT CHAIN 202 286 COAT PROTEIN VP4.  
 FT CHAIN 287 504 COAT PROTEIN VP2.  
 FT CHAIN 505 725 COAT PROTEIN VP3.  
 FT CHAIN 726 937 COAT PROTEIN VP1.  
 FT CHAIN 938 1578 CORE PROTEIN P52.  
 FT CHAIN 1579 1601 GENOME-LINKED PROTEIN VP61.  
 FT CHAIN 1602 1625 GENOME-LINKED PROTEIN VP62.  
 FT CHAIN 1626 1849 GENOME-LINKED PROTEIN VP63.  
 FT CHAIN 1850 1863 PROTEASE P20B.  
 FT CHAIN 1864 2333 RNA-DIRECTED RNA POLYMERASE P56A.  
 FT CHAIN 202 202 N-mvristoyl glycine (in host).  
 FT LIPID 396 396 S -> C (IN REF. 2).  
 FT CONFLICT 396 396 P -> L (IN REF. 2).  
 FT CONFLICT 632 632  
 SQ SEQUENCE 2333 AA; 259645 MW; 4FC667DCC521BC60 CRC64;

Query Match 5.1%; Score 145.5; DB 1; Length 2333;  
 Best Local Similarity 20.8%; Pred. No. 0.02;  
 Matches 92; Conservative 47; Mismatches 203; Indels 101; Gaps 13;

QY 70 NSPGHILNLELGPDLNPNYLAHLSRMNGYAGVEVQVLLAGNAFTAGKILFAAIPPNPL 129  
 DB 358 DKPFGYLTCLBELPTDHHGVFGHLVDSYAYMRNGMDVEVSAGVNGQFNGGCLLVAMVPEWKA 417  
 QY 130 VDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVYHFNNOQPQRMRLVAMLYPLRSN 189  
 DB 418 FDTREKYQLTLFPHQFI-----SPRTNMTAHTVYVLGV 451  
 QY 190 GSGDDVFTVSCRVLTPPTDPEFIVLPSPVESKTKPFTLPTLTISELTNSRFPPIBOL 249  
 DB 452 NRYDQY-----KQHKPWTLLVVMVLSPLTVSNVTAAPQIKV 485  
 QY 250 YT-APNNTNVQCONGRCTLDGELQGTOLLSSAVCFLOGRTVADNGNMDQNLQLTY 307  
 DB 486 YANIAPTIVYHVA-----GELPSKEGIFPVACADGYGGLVTTDPKATDPVYGVKNV 535  
 QY 308 PNGASY-----DPTDEVAPLGTQDFSGMLYGVLTQDNV-----NVSTGEAKNAKG 353  
 DB 536 PKNTYNGRFTNLLDVAEACFTLRFDD--GKPYVVTTRADDTRLAKFDVSLA-AKHMSN 592  
 QY 354 IYVISTTSKFTPKIGSIGLHSI---TEHVHPNQSRFTPVGVAVDVNTFPQOMVLPVHAG 410  
 DB 593 TYLSGIAQYTYQYSTINLHFMFTGSTDASKARYWAYIYPPGVETPDTPPEAAHCHIAEW 652  
 QY 411 SLALNTNLAAPVAPTPGQELLFFRSRVPVCGVQGLQGDADFICLLLPQBWNNHYQEAAPS 470  
 DB 653 DTGLNSKFTFSI-PYVSAADYAYTASDTAETTNVQG-----WVCVYQITHGKA 699  
 QY 471 QADVALLIRYVNDTGRTLFEAKL 493  
 DB 700 ENDTLL-----VSASAGKO-FELRL 718

DR InterPro; IPR007094; RNA\_pol\_PSVir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF05408; Peptidase\_C28; 1.  
 DR Pfam; PF00073; rhv; 3.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PR00918; CALICVIRUSNS.  
 DR PRINTS; PR01542; FMDVP1COAT.  
 DR POLYProtein; Coat protein; Core protein; RNA-directed RNA polymerase;  
 KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;  
 KW Myristate; Lipoprotein.  
 KW CHAIN 1 200 NONSTRUCTURAL PROTEIN P20A.  
 FT CHAIN 201 285 COAT PROTEIN VP4.  
 FT CHAIN 286 503 COAT PROTEIN VP2.  
 FT CHAIN 504 723 COAT PROTEIN VP3.  
 FT CHAIN 724 937 COAT PROTEIN VP1.  
 FT CHAIN 938 953 CORE PROTEIN X.  
 FT CHAIN 954 1107 CORE PROTEIN P4.  
 FT CHAIN 1108 1425 CORE PROTEIN P19.  
 FT CHAIN 1426 1578 CORE PROTEIN P19.  
 FT CHAIN 1579 1601 GENOME-LINKED PROTEIN VPGL.  
 FT CHAIN 1602 1625 GENOME-LINKED PROTEIN VPGL.  
 FT CHAIN 1626 1649 GENOME-LINKED PROTEIN VPGL.  
 FT CHAIN 1650 1862 PROTEASE.  
 FT CHAIN 1863 2332 RNA-DIRECTED RNA POLYMERASE.  
 FT CHAIN 201 201 N-myristoyl glycine (in host).  
 FT LIPID 2332 AA; 259408 MW; EE77DA739CBEDC6A CRC64;  
 SQ SEQUENCE 2332 AA; 259408 MW; EE77DA739CBEDC6A CRC64;  
 Query Match 4.5%; Score 130; DB 1; Length 2332;  
 Best Local Similarity 22.4%; Pred. No. 0.28;  
 Matches 76; Conservative 31; Mismatches 149; Indels 84; Gaps 10;  
 QY 79 LELGPDNLNLYLAHLRMYNGYAGGVQVLLAGNAFTAGKILPAALPPNPLVDUMISPAQI 138  
 Db 366 LELFTDHHGVFGLVDSYAYMRNGWDVEVSANGQNGGCLLVAMVPEWKTFTDREYQL 425  
 QY 139 TMLPHLIVDVTLEPIMTLPDVRNVEYHFNQPPRMRLVAMLYPLRNGSGDDVFTV 198  
 Db 426 TLFPHQFI-----SPRTNTHAITYPLGVNRYDQY--- 456  
 QY 199 SCRVLTRPTDFEYILVPPSVESKTPFTLPILTISELNSRFPPIEOLYT--APNET 256  
 Db 457 -----KKHKPWTLLVMVLSPLTVSNTAATQIKVYANIAITYV 493  
 QY 257 NVVQCNQRCYLDGELQGTQLSSAVCFLOGRV-----ADNGDNWQNLQLQTYPNG 310  
 Db 494 HVA-----GELPSKVGIFPVACSDGYGGLVTTDPKTADPVYKENVPPKTYNPRR 543  
 QY 311 AS--YDPTDEVPAPIGTQDFSGMLYGVLTQDNV-----NVSTGEAKNAKGIYISTSGK 362  
 Db 544 FTNLLDVAEACPTFLCFDD--GKPYVYVTRTDTRLAKFDVSLA-AKHMSNTYLSGIAQY 600  
 QY 363 FTFPKIGSIGLHSI---TEHHPNQOSRTPVGVAVDENTP 399  
 Db 601 YTYSGSTINLHPMTGSTDSKARYVMVAYIPPGVETPPETP 640  
 RESULT 12  
 POLG CX16T  
 ID POLG CX16T STANDARD; PRT; 2193 AA.  
 AC Q9QF31;  
 DT 10-OCT-2003 (Rel. 42, Created)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].  
 DE Cxsackievirus A16 (strain Tainan/5079/98).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Enterovirus.  
 OC NCBI\_TaxID=231417;

DR POLG\_FMDVA STANDARD; PRT; 2332 AA.  
 AC P03308; P03312; Q65038; Q65039; Q65040; Q65041; Q65042; Q65043;  
 AC Q65044; Q65045; Q65046; Q65047;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat proteins VP1 TO VP4; Core proteins X, P14, P41, P19; Genome-linked proteins VPGL TO VPGL3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48)].  
 DE Foot-and-mouth disease virus (strain A12) (Aphthovirus A) (FMDV).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Aphthovirus.  
 OC NCBI\_TaxID=12114;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE=85211015; PubMed=2987518;  
 RX Robertson B.H., Grubman M.J., Weddell G.N., Moore D.M., Welsh J.D.,  
 RA Fischer T., Dowbenko D.J., Yansura D.G., Small B., Kleid D.G.;  
 RT "Nucleotide and amino acid sequence coding for polypeptides of  
 foot-and-mouth disease virus type A12.";  
 RL J. Virol. 54:651-660(1985).  
 [2]  
 RN SEQUENCE OF 1863-2332 FROM N.A.  
 RP MEDLINE=83225613; PubMed=6305004;  
 RX Robertson B.H., Morgan D.O., Moore D.M., Grubman M.J., Card J.,  
 RA Fischer T., Weddell G.N., Dowbenko D.J., Yansura D.G.;  
 RT "Identification of amino acid and nucleotide sequence of the  
 foot-and-mouth disease virus RNA polymerase.";  
 RL Virology 126:614-623(1983).  
 [3]  
 RN SEQUENCE OF 715-955 FROM N.A.  
 RP MEDLINE=82061853; PubMed=6272395;  
 RX Kleid D.G., Yansura D.G., Small B., Dowbenko D.J., Moore D.M.,  
 RA Grubman M.J., McKercher P.D., Morgan D.O., Robertson B.H.,  
 RA Buchrach H.L.;  
 RT "Cloned viral protein vaccine for foot-and-mouth disease: responses  
 in cattle and swine.";  
 RL Science 214:1125-1129(1981).  
 CC -1- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the  
 poliovirus polyprotein. In other picornavirus reactions Glu may be  
 substituted for Gln, and Ser or Thr for Gly.  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 (RNA) (N)  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 VP3, and VP4.  
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 each of which is composed of one copy each of proteins VP1, VP2,  
 VP3, and VP4.  
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.ebi.ac.uk/submit/>  
 or send an email to [license@ebi.ac.uk](mailto:license@ebi.ac.uk)).  
 CC -----  
 CC EMBL; M10975; AAA42593.1; -;  
 DR EMBL; J02187; AAA42670.1; -;  
 DR MEROPS; C03.008; -;  
 DR InterPro; IPR004004; Calici\_pol\_hel.  
 DR InterPro; IPR009003; Cys Ser trypsin.  
 DR InterPro; IPR004080; FMDVP1COAT.  
 DR InterPro; IPR008739; Peptidase\_C28.  
 DR InterPro; IPR001676; Rhv.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.





```

RT Sinorhizobium meliloti pSymA megaplasmid.
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888 (2001).
CC -!- FUNCTION: RECEPTOR FOR THE SIDEROPHORE RHIZOACTIN.
CC -!- SUBCELLULAR LOCATION: Outer membrane.
CC -!- SIMILARITY: Belongs to the tonB-dependent receptor family.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF110737; AAD09419.1; -.
CC DR EMBL; AE007312; AAK65923.1; -.
CC DR PIR; A95420; A95420.
CC DR PIR; T46821; T46821.
CC DR InterPro; IPR000531; TonB boxC.
CC DR Pfam; PF00593; TonB dep Rec; 1.
CC DR PROSITE; PS00430; TONB_DEPENDENT_REC_1; FALSE NEG.
CC DR PROSITE; PS01156; TONB_DEPENDENT_REC_2; FALSE NEG.
CC DR Outer membrane; Iron transport; Transport; TonB box; Signal; Receptor;
CC KW Plasmid; Complete proteome.
CC FT SIGNAL 1 26 POTENTIAL.
CC FT CHAIN 27 746 RHIZOACTIN RECEPTOR.
CC FT SITE 40 47 TONB BOX.
CC FT SITE 729 746 TONB C-TERMINAL BOX.
CC FT SITE 746 AA; 80633 MW; 16A844A025D5B23 CRC64;
CC SEQUENCE 746 AA; 80633 MW; 16A844A025D5B23 CRC64;
CC -----
Query Match 4.3%; Score 123.5; DB 1; Length 746;
Best Local Similarity 19.7%; Pred. No. 0.18;
Matches 110; Conservative 76; Mismatches 184; Indels 189; Gaps 26;
Oy 10 PSNDGAASLVPEGINETMPLEPVAGSI--AAPVAGTNIIDPWIRNFVQAPNGEFTVS 67
Db 91 PASDGARTSGQNLRGPPILVDGVSMNSARSLSRQFDAIDFF-NIERVEVLSGATAIY 149
Oy 68 PRNSPGEL-LNLEGLPDLNPLYLAH---LSRMVNGVAGGV-----EVQ 106
Db 150 GGNATGGIINIITKKGADEPGL-HAEVTGGMGSGFAGSQDFDRAAGAVTYNSENWDAR 208
Oy 107 VLLAGN---AF--TAGKILFAAIPP-----NFLVDMISPAQITWMLPHLIVDVRTLE---- 152
Db 209 LSIAGNRTGAFYDGSGLTLLPIDITQTSTAFNERIDLNG-----SIGQIDDDRVEFSQG 263
Oy 153 -----PIMTLPDVRNVF-----YHFNQNPQPRMLVAMLYT----- 184
Db 264 YFDSKQSDYGLIYCPFFAALAD-PSLPETRSGVSEDFNPQTRRSLNLTNTYTDNDVFGQQ 322
Oy 185 -----PLRNSGSD-----DVFTVSCRLVTRPTDPFEIYLP 217
Db 323 LLLQGSYRTERIKPHFPFASGNSETGPYFGSSQDTDYGIRAAALVAEPTALKITYGID 382
Oy 218 PSVESKT-----KPFTLPILITISLNSRFPPIPIEQLYTAPN 254
Db 383 ADMDSFTARQNFDMVAAGQSGGLDNFTIGKTLGYPISDVSTVAGF-----AEASYEATD 437
Oy 255 ETNVVQCQNGRCTLGELQ-----GTTQLLSAVCFLOGR-TVADNGDNWDQN 301
Db 438 -----RLTLNGGVRYQFNVTEVSDFIGAAQ-----QVAILOGRATSDATTPGGEVN 483
Oy 302 LIQLTYPCASVDPTD-----EVRAP-----LGTQDFSGMLYGLVLTQDNVNS 344
Db 484 YDAALFSAGATVQLTNTQOYVAFNSQGFELPDPKAYGIGNYVFSFGHTLVNSVNWGDS 543
Oy 345 TGEAKNAKGIYISTTSKFTPKIGSTGLHSITEH-VHPNQQRFTPV----- 390
Db 544 ALBAIKTNSFEIGVRLDDGTFTNLEAAYVSLSDRSINLNRSLAVEIIDRRERVYIGSK 603
Oy 391 -GVAVDENTPFQWVLP 408
Db 604 AGVKLDHG--FDVGLVGHW 620

```

```

RESULT 14
POLG HE701 STANDARD; PRT; 2194 AA.
ID POLG HE701 STANDARD; PRT; 2194 AA.
AC P32537;
DT 01-OCT-1993 (Rel. 32, last sequence update)
DT 01-NOV-1995 (Rel. 42, last annotation update)
DE Genome polypeptide [P32537]; Coat protein VP4 (PIA); Coat protein VP2
DE (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Core protein
DE (P2A); Core protein P2B; Core protein P2C; Core protein P3A; Genome-
DE linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48) [1].
OS Human enterovirus 70 (strain J670/71) (EV 70).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=31915;
RN [1]
SEQUENCE FROM N.A.
MEDLINE=91037960; PubMed=2172447;
RX Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D.,
RA Minor P.D., Almond J.W.;
RT "The complete nucleotide sequence of enterovirus type 70:
RT relationships with other members of the picornaviridae.";
RL J. Gen. Virol. 71:2291-2299 (1990).
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC O/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D00820; BAA18891.1; -.
CC DR PIR; A36253; GNNYE7.
CC DR HSP; P03300; IPOV.
CC DR MEROPS; C03.UBA; -.
CC DR MEROPS; C03.UBP; -.
CC DR InterPro; IPR003593; AAA ATPase.
CC DR InterPro; IPR004004; Calici_pol_hel.
CC DR InterPro; IPR009003; Cys_Ser_trypsin.
CC DR InterPro; IPR001199; pept_3C_picorn.
CC DR InterPro; IPR000081; peptidase_C3.
CC DR InterPro; IPR003338; Pico_P1A.
CC DR InterPro; IPR002527; Pico_P2B.
CC DR InterPro; IPR001676; Rhv.
CC DR InterPro; IPR000605; RNA_helicase.
CC DR InterPro; IPR007095; RNA_pol_DS_PS.
CC DR InterPro; IPR001205; RNA_pol_P3D.
CC DR InterPro; IPR007094; RNA_pol_PSVir.
CC DR InterPro; IPR008975; Viral_cap_coat.
CC DR Pfam; PF00548; Cys-protease_3C; 1.
CC DR Pfam; PF02226; Pico_P1A; 1.
CC DR Pfam; PF00947; Pico_P2A; 1.
CC DR Pfam; PF01552; Pico_P2B; 1.
CC DR Pfam; PF00073; rhv; 3.
CC DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC DR Pfam; PF00910; RNA_helicase; 1.
CC DR PRINTS; PR00918; CALICUVIRUSN.

```

```

DR ProDom; PD001125; Cys_protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein. 2 69 COAT PROTEIN VP4.
FT CHAIN 70 319 COAT PROTEIN VP2.
FT CHAIN 320 561 COAT PROTEIN VP3.
FT CHAIN 562 871 COAT PROTEIN VP1.
FT CHAIN 872 1014 CORE PROTEIN P2A.
FT CHAIN 1015 1113 CORE PROTEIN P2B.
FT CHAIN 1114 1443 CORE PROTEIN P2C.
FT CHAIN 1444 1532 CORE PROTEIN P3A.
FT CHAIN 1533 1554 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1555 1737 PICORNAIN 3C.
FT CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 N-myristoyl glycine (in host) (by
FT ACT SITE 1701 1701 PROTEASE (POTENTIAL).
FT ACT SITE 1715 1715 PROTEASE (POTENTIAL).
SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;

Query Match 4.1%; Score 118; DB 1; Length 2194;
Best Local Similarity 21.9%; Pred. No. 2;
Matches 57; Conservative 39; Mismatches 114; Indels 50; Gaps 10;

QY 48 IDPWRTNFVQPNQ-----EFTVSPRNSGTEILLNLEL-----GPDLPYLHLRLMYNG 98
DQ 368 IESMEINNVNDASGVERLVRVQISQSDMDQLFNIPDLIDLEGPRVTLGNISRYTH 427
QY 99 YAGGVEVQVLLAGNAFTAGKILFAIPFNFLVDMISPA---QITMLPHLIVDVRLTLPIM 155
DQ 428 WSGSLEMTFMFGSGFTTGKLIICVTPG---GSSPTDRMQAMLAHVWDFGLQSSIT 483
QY 156 TPLDVRNVFYH-FNNQOPQRMVLYMAYLTPLRN-----GSGDVFVTVSCRV----- 202
DQ 484 IIPWISGSHYRMENTDAKAINANVGYVTCFMQTLNVLAPVGAADQCYIVGMVAKKDFNL 543
QY 203 -LTRPTDPF-----EFYLVPPSVESKTKP--FTLPILTISE--LTNSR 241
DQ 544 RLMRSPDIGQAILPEQAATQIGEVKTVANTVESEIKAEGLVIPSNAVETGATNT 603
QY 242 FPIPIEQLYTAPNETNVQC 261
DQ 604 EPEEALQRTVLNMGHTAEC 623

RESULT 15
POLG_CX16G STANDARD; PRT; 2193 AA.
AC Q65900.
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2
DE (PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Core protein
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
DE linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Coxsackievirus A16 (strain G-10).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=69159;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=94303216; PubMed=8030260;
RA Povty T., Hyypiaie T., Horsnell C., Kinnunen L., Hovi T., Stanway G.;
RT "Molecular analysis of coxsackievirus A16 reveals a new genetic group
RT of enteroviruses.";
RL Virology 202:982-987(1994).
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular

```

```

CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC [RNA](N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; U05876; AAA50478.1; -.
CC HSP; P03300; IPOV.
CC MEROPS; C03.022; -.
CC DR InterPro; IPR003593; AAA_ATPase.
CC DR InterPro; IPR004004; Calici_pol_hel.
CC DR InterPro; IPR009003; Cys_ser_trypsin.
CC DR InterPro; IPR001199; Pept_3C_picorn.
CC DR InterPro; IPR000081; Peptidase_C3.
CC DR InterPro; IPR003138; Pico_P1A.
CC DR InterPro; IPR002527; Pico_P2B.
CC DR InterPro; IPR001676; Rnv.
CC DR InterPro; IPR000605; RNA_helicase.
CC DR InterPro; IPR007095; RNA_pol_DS_P8.
CC DR InterPro; IPR001205; RNA_pol_P3D.
CC DR InterPro; IPR007094; RNA_pol_P5vir.
CC DR InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00548; Cys-protease-3C; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF00073; thv; 3.
CC DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC DR Pfam; PF00910; RNA_helicase; 1.
CC DR PRINTS; PR00918; CALICVIRUSNS.
CC DR ProDom; PD001125; Cys_protease_3C; 1.
CC DR ProDom; PD001306; Pico_P2A; 1.
CC DR ProDom; PD001274; Pico_P2B; 1.
CC DR SMART; SM00382; AAA; 1.
CC DR Polyprotein; Coat protein; Core protein; Transferase;
CC RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
CC Lipoprotein. 2 69 COAT PROTEIN VP4.
CC CHAIN 70 323 COAT PROTEIN VP2.
CC CHAIN 324 565 COAT PROTEIN VP3.
CC CHAIN 566 862 COAT PROTEIN VP1.
CC CHAIN 863 1012 CORE PROTEIN P2A.
CC CHAIN 1013 1111 CORE PROTEIN P2B.
CC CHAIN 1112 1440 CORE PROTEIN P2C.
CC CHAIN 1441 1526 CORE PROTEIN P3A.
CC CHAIN 1527 1548 GENOME-LINKED PROTEIN VPG.
CC CHAIN 1549 1731 PICORNAIN 3C.
CC CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.
CC CHAIN 2 2 N-myristoyl glycine (in host) (by
CC ACT SITE 1695 1695 PROTEASE (POTENTIAL).
CC ACT SITE 1709 1709 PROTEASE (POTENTIAL).
SQ SEQUENCE 2193 AA; 243209 MW; 04B3BC572A76E38 CRC64;

Query Match 4.1%; Score 116.5; DB 1; Length 2193;
Best Local Similarity 20.9%; Pred. No. 2.6;

```

	Matches	86;	Conservative	53;	Mismatches	150;	Indels	123;	Gaps	187;
Qy		1	MKQASNDAA	PNSDGA	SILVPEGINET	-----MP-----	LEPVAGASIAAPVAGQTNIIDP	50		
Db		329	LKPGTNQFLTDDGVSAPILGP	PHPTPAIHIPGEVRNLLEICRVETILEVNVLNQSNETTP	388					
Qy		51	WIRTNF	---VQAPNGEFTVSPRNPGBEILNLBGLPDLPYLAHLSRMYNGVAGGVEVQV	107					
Db		389	MQRLCFPVS	VQSKTGELCAVFRADFG	-----RNGPWQSTILGQLCKRYTQMWSGLEWTF	442				
Qy		108	LLAGNAFTAGKTFLFAAIPPNFLVDMI	SPA-QITWL--PHLIVDVTRLEBPIMTLPDVRNV	164					
Db		443	MFAGSGFMATGKMILLIAYTPGGGV	---PADRLTAMLGTHVINDFGLOSVTLVIPWISNT	498					
Qy		165	FYFHNNQPQPRMRLVAMLYTPLRSNGSGDDVF	TVCRCVLTRPTDPDEFYILVPSPVESKT	224					
Db		499	HYRAH	:	-----AKGYFDY-----	511				
Qy		225	KPTLTPIILTISELTNSRRPIP	-----IEQLYTAPNETNVVQCQNGR	-----CTILDGLQ	273				
Db		512	--YTTGTITITWTQTVNVVPI	GAPTAYIVALAAQDNFTMKLCKDTEDIEQSANIQQD--	567					
Qy		274	GTTQLLSSAVCF	LQGRRTVADNGDNWDONLLQITVPNGASYPDTEVPALGTQDSGMXY	333					
Db		568	GIADMIDQAVTSRVGRALTS	-----LOVE-PTAANTNASEH-----RLGT----	GLVP	610				
Qy		334	GVUTQDNVVNSTGEAKNAKGIYI	-----STTSCKTPKIGSIGL	372					
Db		611	ALQAAE	-----TGASSNAQDENLIETRCLVNLHHSTQETTIGNFFPSRAGLSVI	657					

Search completed: June 1, 2004, 13:47:23  
Job time : 8.54765 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model  
Run on: June 1, 2004, 13:35:27 ; Search time 31.556 Seconds  
(without alignments)  
5419.275 Million cell updates/sec

Title: US-09-926-799-9  
Perfect score: 2863  
Sequence: 1 MKMASNDAPNDGAASLVP.....VNQFSLAPMTGNGRRRIQ 542

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL\_25.\*

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phase.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_rvirus.\*
- 16: sp\_bacteriaph.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2863	100.0	542	12	Q8JW45
2	2822	98.6	542	12	Q8JMP1
3	2049.5	71.6	535	12	Q91115
4	2048.5	71.6	535	12	Q9QMK6
5	2047.5	71.5	535	12	Q917X7
6	2046.5	71.5	535	12	Q915C6
7	2028.5	70.9	535	12	Q8V784
8	2020.5	70.6	535	12	Q88104
9	2016	70.4	548	12	Q915C7
10	2015.5	70.4	535	12	Q917Y9
11	2010.5	70.2	535	12	Q917Y3
12	2007.5	70.1	535	12	Q917X4
13	2006.5	70.1	535	12	Q916E4
14	2003.5	70.0	535	12	Q917Y0
15	2002.5	69.9	535	12	Q916E6
16	1994.5	69.7	535	12	Q91V39

17	1988.5	69.5	535	12	Q91V51	Q91V51 human calic
18	1980	69.2	542	12	Q66293	Q66293 calicivirus
19	1973	68.9	542	12	O12305	O12305 snow mounta
20	1973	68.9	542	12	Q8ORD6	Q8ORD6 snow mounta
21	1970	68.8	542	12	Q96877	Q96877 snow mounta
22	1966	68.7	542	12	Q913B6	Q913B6 human calic
23	1943	67.9	548	12	Q66296	Q66296 calicivirus
24	1942	67.8	548	12	Q8V0P5	Q8V0P5 human calic
25	1942	67.8	548	12	Q8V0P4	Q8V0P4 human calic
26	1935.5	67.6	547	12	Q8V775	Q8V775 norwalk vir
27	1934.5	67.6	547	12	Q918A7	Q918A7 norwalk-vir
28	1933.5	67.5	547	12	Q8V773	Q8V773 norwalk-vir
29	1928	67.3	540	12	Q917X1	Q917X1 norwalk-lik
30	1928	67.3	548	12	Q8V768	Q8V768 norwalk vir
31	1928	67.3	548	12	Q918A1	Q918A1 norwalk-lik
32	1928	67.3	548	12	Q8BC87	Q8BC87 human calic
33	1924	67.2	540	12	Q917W8	Q917W8 norwalk-lik
34	1923	67.2	548	12	Q916E5	Q916E5 human calic
35	1922	67.1	548	12	Q83880	Q83880 norwalk vir
36	1921	67.1	548	12	Q88291	Q88291 small round
37	1921	67.1	548	12	Q80FL1	Q80FL1 human calic
38	1920	67.1	548	12	Q80FK8	Q80FK8 human calic
39	1920	67.1	548	12	Q80FK5	Q80FK5 human calic
40	1919	67.0	548	12	Q80FK9	Q80FK9 human calic
41	1918	67.0	538	12	Q91165	Q91165 human calic
42	1918	67.0	548	12	Q91725	Q91725 norwalk-lik
43	1916	66.9	548	12	Q9PYA7	Q9PYA7 human calic
44	1916	66.9	548	12	Q8BC90	Q8BC90 human calic
45	1916	66.9	548	12	Q80FK4	Q80FK4 human calic

ALIGNMENTS

RESULT 1  
Q8JW45 PRELIMINARY; PRT; 542 AA.  
ID Q8JW45  
AC Q8JW45;  
DC 01-OCT-2002 (TremBLrel. 22, Created)  
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)  
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)  
DE Capsid protein.  
GN CAPSID.  
OS Norwalk-like virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=95340;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Kashiwa47;  
RA Natori K., Takeda N.;  
RT "Genetic and antigenic relationship among Norwalk-like viruses.";  
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB078334; BAC05515.1; -  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 542 AA; 59149 MW; 5C28719C65F6C667 CRC64;  
Query Match 100.0%; Score 2863; DB 12; Length 542;  
Best Local Similarity 100.0%; Pred. No. 6.9e-217; Mismatches 0; Indels 0; Gaps 0;  
Matches 542; Conservative 0;

QY 1 MKMASNDAPNDGAASLVPGEINETMPLEPVAGASIAAPVAGQTNIIDPRTNFVQAP 60  
Db 1 MKMASNDAPNDGAASLVPGEINETMPLEPVAGASIAAPVAGQTNIIDPRTNFVQAP 60  
QY 61 NGEFTVSPRNSPGEILLNLELGPDLNLYLAHLRMYNGYAGGVEVQVLLAGNAFTAGKIL 120

```

Db 61 NGEFTVSPRNSPGEILLNLELGPDLNPNYLAHLRSMYNGYAGGVEQVLLAGNAFTAGKIL 120
121 FFAAIPPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPOPRMELVA 180
121 FFAAIPPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPOPRMELVA 180
181 MLYTPLRNSGDDVFTVSCRVLTPTDFFIYLVPPSVESKTKPFTLPILTISELTNS 240
181 MLYTPLRNSGDDVFTVSCRVLTPTDFFIYLVPPSVESKTKPFTLPILTISELTNS 240
241 RPIPIEQIYATPNETNVCQNGRCTLGDELQGTTLQSSAVCFLOGRTVADRGDNWDQ 300
241 RPIPIEQIYATPNETNVCQNGRCTLGDELQGTTLQSSAVCFLOGRTVADRGDNWDQ 300
301 NLLQLTYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKIYISTTS 360
301 NLLQLTYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKIYISTTS 360
361 GKFTPKIGSIGLHSITEHVHPNQSRFTPVGVAVDENTPFQQWVLPHYAGSLALNTNLAP 420
361 GKFTPKIGSIGLHSITEHVHPNQSRFTPVGVAVDENTPFQQWVLPHYAGSLALNTNLAP 420
421 AVAPTPFGQLLFFRSRVPVCVQLOGQDAFIDCLLPQEWNVHFIYQEAAPSQADVALIRYV 480
421 AVAPTPFGQLLFFRSRVPVCVQLOGQDAFIDCLLPQEWNVHFIYQEAAPSQADVALIRYV 480
481 NPDGTGRTLFEAKLHRSGLFIVTSHTGAYPLVVPNGHFRFDSWVNOFYSLAPMGTNGRRR 540
481 NPDGTGRTLFEAKLHRSGLFIVTSHTGAYPLVVPNGHFRFDSWVNOFYSLAPMGTNGRRR 540
541 IQ 542
541 IQ 542

RESULT 2
Q8JMP1 PRELIMINARY; PRT; 542 AA.
AC Q8JMP1;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Major capsid protein VP1.
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]_TaxID=95340;
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Fayetteville/1998/US;
RA Vinje J., Tseng F., Moe C.L.;
RT "Norwalk-like virus strain Hu/NLV/Fayetteville/1998/US, complete ORF2 gene.";
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY113106; RAMS6034.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 542 AA; 59147 MW; 528911EB227893A7 CRC64;

Query Match 98.6%; Score 2822; DB 12; Length 542;
Best Local Similarity 98.5%; Pred. No. 1.2e-213;
Matches 534; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

QY 1 MKMASNDAAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNIIDPWIRNFVQAP 60
Db 1 MKMASNDAAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNIIDPWIRNFVQAP 60
61 NGEFTVSPRNSPGEILLNLELGPDLNPNYLAHLRSMYNGYAGGVEQVLLAGNAFTAGKIL 120
QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPNYLAHLRSMYNGYAGGVEQVLLAGNAFTAGKIL 120
Db 61 NGEFTVSPRNSPGEILLNLELGPDLNPNYLAHLRSMYNGYAGGVEQVLLAGNAFTAGKIL 120
QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPNYLAHLRSMYNGYAGGVEQVLLAGNAFTAGKIL 120
Db 61 NGEFTVSPRNSPGEILLNLELGPDLNPNYLAHLRSMYNGYAGGVEQVLLAGNAFTAGKIL 120

```

```

QY 121 FFAAIPPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPOPRMELVA 180
Db 121 FFAAIPPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPOPRMELVA 180
181 MLYTPLRNSGDDVFTVSCRVLTPTDFFIYLVPPSVESKTKPFTLPILTISELTNS 240
181 MLYTPLRNSGDDVFTVSCRVLTPTDFFIYLVPPSVESKTKPFTLPILTISELTNS 240
241 RPIPIEQIYATPNETNVCQNGRCTLGDELQGTTLQSSAVCFLOGRTVADRGDNWDQ 300
241 RPIPIEQIYATPNETNVCQNGRCTLGDELQGTTLQSSAVCFLOGRTVADRGDNWDQ 300
301 NLLQLTYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKIYISTTS 360
301 NLLQLTYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKIYISTTS 360
361 GKFTPKIGSIGLHSITEHVHPNQSRFTPVGVAVDENTPFQQWVLPHYAGSLALNTNLAP 420
361 GKFTPKIGSIGLHSITEHVHPNQSRFTPVGVAVDENTPFQQWVLPHYAGSLALNTNLAP 420
421 AVAPTPFGQLLFFRSRVPVCVQLOGQDAFIDCLLPQEWNVHFIYQEAAPSQADVALIRYV 480
421 AVAPTPFGQLLFFRSRVPVCVQLOGQDAFIDCLLPQEWNVHFIYQEAAPSQADVALIRYV 480
481 NPDGTGRTLFEAKLHRSGLFIVTSHTGAYPLVVPNGHFRFDSWVNOFYSLAPMGTNGRRR 540
481 NPDGTGRTLFEAKLHRSGLFIVTSHTGAYPLVVPNGHFRFDSWVNOFYSLAPMGTNGRRR 540
541 IQ 542
541 IQ 542

RESULT 3
Q91115 PRELIMINARY; PRT; 535 AA.
AC Q91115;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Schwerin 003/00/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=168852;
RN [1]_TaxID=168852;
RP SEQUENCE FROM N.A.
RA Kuenkel U., Schreier E.;
RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF397905; AAK92128.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58520 MW; 2E97BBD3F3D5FCC1 CRC64;

Query Match 71.6%; Score 2049.5; DB 12; Length 535;
Best Local Similarity 71.7%; Pred. No. 8.1e-153;
Matches 392; Conservative 65; Mismatches 73; Indels 17; Gaps 7;

QY 1 MKMASNDAAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNIIDPWIRNFVQAP 60
Db 1 MKMASNDAAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNIIDPWIRNFVQAP 60
61 NGEFTVSPRNSPGEILLNLELGPDLNPNYLAHLRSMYNGYAGGVEQVLLAGNAFTAGKIL 120
QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPNYLAHLRSMYNGYAGGVEQVLLAGNAFTAGKIL 120
Db 61 NGEFTVSPRNSPGEILLNLELGPDLNPNYLAHLRSMYNGYAGGVEQVLLAGNAFTAGKIL 120
QY 121 FFAAIPPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPOPRMELVA 180
Db 121 FFAAIPPNFLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFYHFNQPOPRMELVA 180

```

```
QY 181 MLYTLRNSGDDVFTVSCRVLTPTDPTDFBFIYLVPPSVESKTKPFTLPILTISLNTS 240
DB 181 MLYTLRNSGDDVFTVSCRVLTPTDPTDFBFIYLVPPSVESKTKPFTLPILTISLNTS 240
QY 241 RPIPIEOLYTAPNETNVVQCONGRCTLDGELQGTTLQLLSSAVCFLOGR---TVADNGDN 297
DB 241 RFPVIDELYTSPNESLVVQNGRCALDGLQGTTLQLLPTAICSFGRINQKVSGENHV 300
QY 298 WDQNLQLTYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKIYI 356
DB 301 WN---MQVTNINGTPTDPTDVPAPLGTDFSGKLFGLVLSQRDHN-----ACRSHDAVI 352
QY 357 STTSCKFTPKIGSLGHSITE-IHVPNQOSRFTPVGVAVDENTPQQWVLPHYAGSLAIN 415
DB 353 ATNSAKFTPKLGAIQIGTWEEDDVHINOPTKFTPVGLF--ENEGFNQWTLPNYSGALTIN 410
QY 416 TNLPAVAPTFFGEQLLFFRSRVPVCVQGLQGDADFDCLLPQEWVNHFFVQEAAPSQADVA 475
DB 411 MGLAPPVAPTFFGEQILFFRSHIPLKGVV--ADPVIDCLLPQEWIQLHYQESAPSQSDVA 468
QY 476 LIRVYNPDTGRTLFEAKLHRSFTVSHGTGAYPLVVPNGHFRPDSWVNFYSLAPMGTG 535
DB 469 LIRFTNPDTGRVLFPEAKLHRSYITVANTGSRPIVVPANGYFRPDSWVNFYSLAPMGTG 528
QY 536 NGRRIQ 542
DB 529 NGRREVQ 535
```

## RESULT 4

```
Q9QMK6
ID Q9QMK6 PRELIMINARY; PRT; 535 AA.
AC Q9QMK6;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-MAY-2000 (TremBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Chitta virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=104901;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Aichi76-96;
RX MEDLINE=20471382; PubMed=11021399;
RA Kobayashi S., Sakae K., Suzuki Y., Iehiko H., Kamata K., Suzuki K.,
RA Natori K., Miyamura T., Takeda N.;
RT "Expression of recombinant capsid proteins of chitta virus, a
RT genogroup II Norwalk virus, and development of an ELISA to detect the
RT viral antigen.";
RL Microbiol. Immunol. 44:687-693(2000).
DR EMBL; AB032758; BAA84716.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58504 MW; 5CCEBC91BFC7B8B CRC64;
```

```
Query Match 71.6%; Score 2048.5; DB 12; Length 535;
Best Local Similarity 71.7%; Pred. No. 9.8e-153;
Matches 392; Conservative 64; Mismatches 74; Indels 17; Gaps 7;
QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNTIIDPWIRTNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEANNTMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60
QY 61 NGEFTVSPNSGCEILLNLEIGDPLNLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120
DB 61 NGEFTVSPNSGCEILLNLEIGDPLNLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120
QY 121 FAALPNEFLVDMISPAQITMLPHLVDVRLTLEPINTPLPDVNFVYHFNQPPMRLLVA 180
DB 121 FAAPVPHPLENISPGQITMFFHVIDVRLTLEPINTPLPDVNFVYHFNQPPMRLLVA 180
```

```
QY 181 MLYTLRNSGDDVFTVSCRVLTPTDPTDFBFIYLVPPSVESKTKPFTLPILTISLNTS 240
DB 181 MLYTLRNSGDDVFTVSCRVLTPTDPTDFBFIYLVPPSVESKTKPFTLPILTISLNTS 240
QY 241 RPIPIEOLYTAPNETNVVQCONGRCTLDGELQGTTLQLLSSAVCFLOGR---TVADNGDN 297
DB 241 RFPVIDELYTSPNESLVVQNGRCALDGLQGTTLQLLPTAICSFGRINQKVSGENHV 300
QY 298 WDQNLQLTYPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKIYI 356
DB 301 WN---MQVTNINGTPTDPTDVPAPLGTDFSGKLFGLVLSQRDHN-----ACRSHDAVI 352
QY 357 STTSCKFTPKIGSLGHSITE-IHVPNQOSRFTPVGVAVDENTPQQWVLPHYAGSLAIN 415
DB 353 ATNSAKFTPKLGAIQIGTWEEDDVHINOPTKFTPVGLF--ENEGFNQWTLPNYSGALTIN 410
QY 416 TNLPAVAPTFFGEQLLFFRSRVPVCVQGLQGDADFDCLLPQEWVNHFFVQEAAPSQADVA 475
DB 411 MGLAPPVAPTFFGEQILFFRSHIPLKGVV--ADPVIDCLLPQEWIQLHYQESAPSQSDVA 468
QY 476 LIRVYNPDTGRTLFEAKLHRSFTVSHGTGAYPLVVPNGHFRPDSWVNFYSLAPMGTG 535
DB 469 LIRFTNPDTGRVLFPEAKLHRSYITVANTGSRPIVVPANGYFRPDSWVNFYSLAPMGTG 528
QY 536 NGRRIQ 542
DB 529 NGRREVQ 535
RESULT 5
Q917X7
ID Q917X7 PRELIMINARY; PRT; 535 AA.
AC Q917X7;
DT 01-DEC-2001 (TremBLrel. 19, Created)
DT 01-DEC-2001 (TremBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS ORF2.
OS Norwalk-like virus NLV/Honolulu/314/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171852;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Honolulu/314/1994/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Honolulu/314/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Honolulu/314/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Honolulu/314/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
```

```

DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1. D839C359F86DC1F9 CRC64;
SQ SEQUENCE 535 AA; 58505 MW; 58505 MW; 58505 MW;

Query Match 71.5%; Score 2046.5; DB 12; Length 535;
Best Local Similarity 71.7%; Pred. No. 1.4e-152;
Matches 392; Conservative 64; Mismatches 74; Indels 17; Gaps 7;

QY 1 MKMAANDAPSDNDGAASLVPEGINETMPLEPVAGASIAAPVAGQTNIIIDPWIRNFVQAP 60
Db 1 MKMAASDAAPSDNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNVIDPWIRNFVQAP 60
QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKIL 120
Db 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVNGYAGGVEQVLLAGNAFTAGKIL 120
QY 121 FFAIPNFVLVDMISPAQITMLPHLIVDRTLEPIMTLPDVRNVPYHFNNOQPRMRLVA 180
Db 121 FFAAVPHFFPLENISPGQIAMFPHVVIDVLTLEPVLPLPDVRNNEFHYNQNEPRMRLVA 180
QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDFEFYIYLVPPSVESKTKPFTLPILTISELTNS 240
Db 181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDFEFYIYLVPPSVESKTKPFTLPILTISELTNS 240
QY 241 RFPPIEQLYTAPNETNVOQNGRCCTLDGELQGTTLSSAVCFLOGR---TVADNGDN 297
Db 241 RFPAPIDELYTSNESLVVQPNQRCALDGELOGTTLPTAICSPRGRINQKVSGENHV 300
QY 298 WDQNLQLQTYPNCASVDPTDEVPAPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKGIYI 356
Db 301 WN---MQVTNNGTPPTDPTEDVPAPLGTDFSGRLFGVLQSQRDHDN-----ACRSHDAVI 352
QY 357 STTSKGKTPKIGSLHSITE-HVHPNQOSRFTPGVAVDENTPFOQWVLPYHAGSLALN 415
Db 353 ATNSAKFTPKLGAIQGTWEEDDVHINQPTKFTPVGLF--EHGDFNQWTLPNYSGALTIN 410
QY 416 TNLAPAVAPTPPGQQLFFRSRVPVCGVQGLQGDADIDCLLPQBWNVHVFQEAAPSQADVA 475
Db 411 MGLAPPVAPTPPGQQLFFRSRHPILKGGV--ADPVIDCLLPQEWIQLHYESAPSQSDVA 468
QY 476 LIRVYNPDGTGRTLEAKLHRSFGFTVSHGTGAYPLVPPNGHFRFDSWVNOFYSLAPMGTC 535
Db 469 LIRFTNPDGTGRTLEAKLHRSFGFTVSHGTGAYPLVPPNGHFRFDSWVNOFYSLAPMGTC 528
QY 536 NGRRRRIQ 542
Db 529 NGRRRRVQ 535

RESULT 7
Q8V784 PRELIMINARY; PRT; 535 AA.
AC Q8V784;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF2 protein (Capaid).
OS Norwalk virus, and
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983, 95340;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk virus; STRAIN=UIGII;
RA Kojima S., Kageyama T., Fukushi S., Hoshino F.B., Shinohara M.,
RA Uchida K., Katori K., Takeda N., Katayama K.;
RT "Genogroup-specific primers for detect Norwalk like virus.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama UI;
RA Katayama K.;
```



Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

```

RL  SPECIES=Norwalk-like virus; STRAIN=Saitama UI.;
RN  SEQUENCE FROM N.A.
RC  MEDLINE=22192455; PubMed=12202225;
RA  Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA  Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,
RA  Gotohori T., Takeda N.;
RT  "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
RT  Viruses.";
RL  Virology 299:225-239 (2002).
DR  EMBL; AB067536; BAB84137.1; -
DR  EMBL; AB039775; BAC11816.1; -
DR  InterPro; IPR004005; Calici_coat.
DR  InterPro; IPR008975; Viral_Cap_coat.
DR  Pfam; PF00915; Calici_coat; 1.
SQ  SEQUENCE 535 AA; 58615 MW; 6B0B2D3162436787 CRC64;

```

```

Query Match 70.9%; Score 2028.5; DB 12; Length 535;
Best Local Similarity 71.1%; Pred. No. 3.7e-151;
Matches 389; Conservative 66; Mismatches 75; Indels 17; Gaps 7;

QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGOTNIIDPWIRTNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEANNETMALEPVAGASIAAPLTGQNNIIDPWIRLNFVQAP 60

QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120
DB 61 NGEFTVSPRNSPGEVLLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKLV 120

QY 121 FAATPPNPLVDMSIPAQITMLPHLIVDRTLEPIMTLPDVRNVFHYHNNOPQPRMLVA 180
DB 121 FAATPPNPLVDMSIPAQITMLPHLIVDRTLEPIMTLPDVRNVFHYHNNOPQPRMLVA 180

QY 181 MLYTPLRNSGDDVFTVSCRVLTPDPEFIYLVPPSVESKTKPFTLPILTISELTNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTPDPEFIYLVPPSVESKTKPFTLPILTISELTNS 240

QY 241 RPPIDIEQLYTAPNETNVVQNGRCITLDGELQGTQLLSSAVCFLOGR---TVADNGDN 297
DB 241 RPPVIDELYTSPNESLVVQNGRCALDGLRGITQLLPTAICFRGRINQKVSNGHV 300

QY 298 WDQNLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKGIYI 356
DB 301 WN---MQVTNINGTGFDETPDVPALGTDFSGKLGVLQRDHDN-----ACRSHDAVI 352

QY 357 STTSCKFTPKIGSIGLHSITE-HVHPNQSRFTPVGVADENTPQQWVLPHYAGSLAIN 415
DB 353 ATNSAKFTPKLGAIQIGTWEEEDVHINQFTKTPVGLF--ENEGFNQWTLPNYSGALTIN 410

QY 416 TNLAPAVAPTPFGEQLLFFRSRVPVCVQGLQGDADFIDCLLPQEWVNFHYQAAAPSOADVA 475
DB 411 MGLAPVAPTPFGEKILFRSHIPLKGV--ADPVIDCLLQEWIQLHYQESAPQSQSDVA 468

QY 476 LIRYVNPDTGRTLPEAKLHRSGLFTVSHGTAYPLVVPNGHFRFDSWVNFYSLAPMGTG 535
DB 469 LIRFTNPDTGRTLPEAKLHRSGLFTVANTGSRPIVVPANGYFRFDSWVNFYSLAPMGTG 528

QY 536 NGRRIQ 542
DB 529 NGRRRVQ 535

```

RESULT 8

```

Q68104 ID Q68104 PRELIMINARY; PRT; 535 AA.
AC Q68104;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-MAR-2001 (TReMBLrel. 16, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Hawaii calicivirus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;

```

```

OC Norovirus.
OX NCBI_TaxID=33750;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RX MEDLINE=943358485; PubMed=8077710;
RA Lew J.F., Kapikian A.Z., Valdesuso J., Green K.Y.;
RT "Molecular characterization of Hawaii virus and other Norwalk-like
RT viruses: evidence for genetic polymorphism among human
RT caliciviruses.";
RL J. Infect. Dis. 170:535-542 (1994).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RA Lew J.F.;
RL Submitted (MAR-1994) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RA Green K.Y., Sosnovtseva S.;
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Hawaii virus/1971/US;
RA Pietneva M.A., Sosnovtseva S., Green K.Y.;
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U07611; AAB97768.2; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58741 MW; DCA252A21084599 CRC64;

```

```

Query Match 70.6%; Score 2020.5; DB 12; Length 535;
Best Local Similarity 71.9%; Pred. No. 1.6e-150;
Matches 391; Conservative 59; Mismatches 83; Indels 11; Gaps 6;

QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGOTNIIDPWIRTNFVQAP 60
DB 1 MKMASNDAAPSNDGAAGLVPEVNNETMALEPVAGASIAAPLTGQNNIIDPWIRNFVQAP 60

QY 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120
DB 61 NGEFTVSPRNSPGEILLNLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKLV 120

QY 121 FAATPPNPLVDMSIPAQITMLPHLIVDRTLEPIMTLPDVRNVFHYHNNOPQPRMLVA 180
DB 121 FAATPPNPLVDMSIPAQITMLPHLIVDRTLEPIMTLPDVRNVFHYHNNOPQPRMLVA 180

QY 181 MLYTPLRNSGDDVFTVSCRVLTPDPEFIYLVPPSVESKTKPFTLPILTISELTNS 240
DB 181 MLYTPLRNSGDDVFTVSCRVLTPDPEFIYLVPPSVESKTKPFTLPILTISELTNS 240

QY 241 RPPIDIEQLYTAPNETNVVQNGRCITLDGELQGTQLLSSAVCFLOGR---TVADNGDN 300
DB 241 RPPVIDELYTSPNESLVVQNGRCITLDGELQGTQLLSSAVCFLOGR---TVADNGDN 300

QY 301 NLLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKGIYISTTS 360
DB 301 NLLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGVLTQ--DNVNVSTGEAKNAKGIYISTTS 360

QY 361 LIRYVNPDTGRTLPEAKLHRSGLFTVSHGTAYPLVVPNGHFRFDSWVNFYSLAPMGTGNGR 538
DB 476 LIRFTNPDTGRTLPEAKLHRSGLFTVANTGSRPIVVPANGYFRFDSWVNFYSLAPMGTGNGR 531

QY 539 RRIQ 542

```

Db 532 RVQ 535

RESULT 9

Q915C7 PRELIMINARY; PRT; 548 AA.

AC Q915C7; (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Capsid protein.

OS Human calicivirus NLV/Erfurt/546/00/DB.

OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.

OC NCBI\_TaxID=173929;

OX [1]

SEQUENCE FROM N.A.

RC STRAIN=NLV/Erfurt/546/00/DE;

RA Kuenkel U., Hoehne M., Schraier E.;

RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF427118; AAL18874.1; -

DR InterPro; IPR004005; Calici\_coat.

DR InterPro; IPR008975; Viral\_cap\_coat.

DR Pfam; PF00915; Calici\_coat; 1.

SQ SEQUENCE 548 AA; 59874 MW; 9A1D8BEE58ADE7D CRC64;

Query Match 70.4%; Score 2016; DB 12; Length 548;

Best Local Similarity 68.9%; Pred. No. 3.7e-150;

Matches 381; Conservative 73; Mismatches 83; Indels 16; Gaps 8;

QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLPVPVAGASIAAPVAGQNIIDPWIRTFVQAP 60

DB 1 MKMASNDAAPSNDGAAGLVPEESNNEVMALEPVVAGASIAAPVAGQNIIDPWIRTFVQAP 60

QY 61 NGFTVSPRNSPGEIILNLELGPDLNPLYLAHLSRMVNGYAGGVQVLLAGNAFTAGKIL 120

DB 61 NGFTVSPRNSPGEIILNLELGPDLNPLYLAHLSRMVNGYAGGVQVLLAGNAFTAGKIL 120

QY 121 FAATPPNPLVDMLSPAQITMLPHLIVDRTLEPIMTLPDVRNVFTHFNQPPMRLLVA 180

DB 121 FAATPPNPLVDMLSPAQITMLPHLIVDRTLEPIMTLPDVRNVFTHFNQPPMRLLVA 180

QY 181 MLYTFLRNSGDDVFTVSCRVLTPTDPDEFIYLVPPSVESKTPFTPLITISLTNS 240

DB 181 MLYTFLRNSGDDVFTVSCRVLTPTDPDEFIYLVPPSVESKTPFTPLITISLTNS 240

QY 241 RPIPIEOLYAPNETNVVQNGRCCTLDGELQGTQTLSSAVCFLOGRT---VAD--NG 295

DB 241 RPIPIEOLYAPNETNVVQNGRCCTLDGELQGTQTLSSAVCFLOGRT---VAD--NG 295

QY 296 DNWDQNLQLYTPNGASVDPDTPVAPLCTODFGMLYGVLTQDNVNVSTGE---AKNA 351

DB 296 DNWDQNLQLYTPNGASVDPDTPVAPLCTODFGMLYGVLTQDNVNVSTGE---AKNA 351

QY 301 THWNVTNVL---NGTDPDPTEDVPAPLGTDFSGIYGVISQRNTVTPGEGDLPANRA 357

DB 301 THWNVTNVL---NGTDPDPTEDVPAPLGTDFSGIYGVISQRNTVTPGEGDLPANRA 357

QY 352 KGIYSTSGKTPKIGSIGLSH-ITEVHPNQOSRFTPVGVA-VDENTPFOQWLPHYA 409

DB 352 KGIYSTSGKTPKIGSIGLSH-ITEVHPNQOSRFTPVGVA-VDENTPFOQWLPHYA 409

QY 358 HEAVIATYSPKTPKLGNIQFSTWETQDVSSQPTKFTPVGLASVDANSHFDQWTLPSYS 417

DB 358 HEAVIATYSPKTPKLGNIQFSTWETQDVSSQPTKFTPVGLASVDANSHFDQWTLPSYS 417

QY 410 GSLALNTNLA PAVATPFGQQLLFRSRVPCVQLOGODAFIDCLLPQEWNVHYQEAAP 469

DB 410 GSLALNTNLA PAVATPFGQQLLFRSRVPCVQLOGODAFIDCLLPQEWNVHYQEAAP 469

QY 418 GALTNNLNLAPSAPVFPCECLLFRSFIP-LKGGYGNA-IDCLMPQEWVQHLQESAP 475

DB 418 GALTNNLNLAPSAPVFPCECLLFRSFIP-LKGGYGNA-IDCLMPQEWVQHLQESAP 475

QY 470 SQADVALIRVNPDTGRTLFEAKLHRSRGTIVSHTGAYPLVVPNGHFRFDSWVNYSL 529

DB 470 SQADVALIRVNPDTGRTLFEAKLHRSRGTIVSHTGAYPLVVPNGHFRFDSWVNYSL 529

QY 476 SLSDVALIRVNPETGRTLFEAKLHENGFLITVARNAGVPVAPTNGYFRFDSWVNYFYL 535

DB 476 SLSDVALIRVNPETGRTLFEAKLHENGFLITVARNAGVPVAPTNGYFRFDSWVNYFYL 535

QY 530 APMTGNGRRRIQ 542

DB 530 APMTGNGRRRIQ 542

QY 536 APWNGSGRRRMQ 548

DB 536 APWNGSGRRRMQ 548

RESULT 10

Q917Y9 PRELIMINARY; PRT; 535 AA.

AC Q917Y9; (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Capsid protein.

OS Norwalk-like virus NLV/Miami/81/1986/US.

OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.

OC NCBI\_TaxID=171848;

OX [1]

SEQUENCE FROM N.A.

RC STRAIN=Hu/NLV/Miami/81/1986/US;

RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;

RT "A one-tube method of reverse transcription-PCR to efficiently amplify a 3-kilobase region from the RNA polymerase gene to the poly(A) tail of small round-structured viruses (Norwalk-like viruses).";

RL J. Clin. Microbiol. 35:570-577(1997).

RN [2]

SEQUENCE FROM N.A.

RC STRAIN=Hu/NLV/Miami/81/1986/US;

RA Ando T., Noel J.S., Fankhauser R.L.;

RT "Genetic classification of 'Norwalk-like viruses.'";

RL J. Infect. Dis. 181:S336-S348(2000).

RN [3]

SEQUENCE FROM N.A.

RC STRAIN=Hu/NLV/Miami/81/1986/US;

RA Ando T., Seto Y., Noel J.S., Monroe S.S., Fankhauser R.L.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF414416; AAL13001.1; -

DR InterPro; IPR004005; Calici\_coat.

DR InterPro; IPR008975; Viral\_cap\_coat.

DR Pfam; PF00915; Calici\_coat; 1.

SQ SEQUENCE 535 AA; 56745 MW; D43CDF7BD8C7225A CRC64;

Query Match 70.4%; Score 2015.5; DB 12; Length 535;

Best Local Similarity 71.8%; Pred. No. 3.9e-150;

Matches 390; Conservative 61; Mismatches 83; Indels 9; Gaps 6;

QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLPVPVAGASIAAPVAGQNIIDPWIRTFVQAP 60

DB 1 MKMASNDAAPSNDGAAGLVPEVNEETMALEPVVAGASIAAPVAGQNIIDPWIRTFVQAP 60

QY 61 NGFTVSPRNSPGEIILNLELGPDLNPLYLAHLSRMVNGYAGGVQVLLAGNAFTAGKIL 120

DB 61 NGFTVSPRNSPGEIILNLELGPDLNPLYLAHLSRMVNGYAGGVQVLLAGNAFTAGKIL 120

QY 121 FAATPPNPLVDMLSPAQITMLPHLIVDRTLEPIMTLPDVRNVFTHFNQPPMRLLVA 180

DB 121 FAATPPNPLVDMLSPAQITMLPHLIVDRTLEPIMTLPDVRNVFTHFNQPPMRLLVA 180

QY 181 MLYTFLRNSGDDVFTVSCRVLTPTDPDEFIYLVPPSVESKTPFTPLITISLTNS 240

DB 181 MLYTFLRNSGDDVFTVSCRVLTPTDPDEFIYLVPPSVESKTPFTPLITISLTNS 240

QY 241 RPIPIEOLYAPNETNVVQNGRCCTLDGELQGTQTLSSAVCFLOGRTVADNGDNWDQ 300

DB 241 RPIPIEOLYAPNETNVVQNGRCCTLDGELQGTQTLSSAVCFLOGRTVADNGDNWDQ 300

QY 241 RFPVPIDELATYSPNEGLVQVQNGRSTLDGELGTTQLVPSNICS LRGRINAHLSNQHR 300

DB 241 RFPVPIDELATYSPNEGLVQVQNGRSTLDGELGTTQLVPSNICS LRGRINAHLSNQHR 300

QY 301 NLLQLYTPNGASVDPDTPVAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAGIYISTTS 360

DB 301 NLLQLYTPNGASVDPDTPVAPLGTQDFSGMLYGVLTQDNVNVSTGEAKNAGIYISTTS 360

QY 301 WNVQTNANGTPDPTEDVPAPLGTDFLANIYGVTSQRNPD-NTCRAHD--GI-LATWS 356

DB 301 WNVQTNANGTPDPTEDVPAPLGTDFLANIYGVTSQRNPD-NTCRAHD--GI-LATWS 356

QY 361 GKFTPKIGSIGLHSEITEH-VHPNQOSRFTPVGVAVDENTPFOQWLPVHAGSLALNTNLA 419

DB 361 GKFTPKIGSIGLHSEITEH-VHPNQOSRFTPVGVAVDENTPFOQWLPVHAGSLALNTNLA 419

QY 357 PKFTPKIGSVLGTWEDRDFDQNPTRFTPTPVGLY--DTDFNQWLPVYSGALTNNMLA 414

DB 357 PKFTPKIGSVLGTWEDRDFDQNPTRFTPTPVGLY--DTDFNQWLPVYSGALTNNMLA 414

Db 532 RVQ 535

RESULT 9

Q915C7 PRELIMINARY; PRT; 548 AA.

AC Q915C7; (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Capsid protein.

OS Human calicivirus NLV/Erfurt/546/00/DB.

OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.

OC NCBI\_TaxID=173929;

OX [1]

SEQUENCE FROM N.A.

RC STRAIN=NLV/Erfurt/546/00/DE;

RA Kuenkel U., Hoehne M., Schraier E.;

RT "Molecular epidemiology of outbreaks of gastroenteritis associated with Norwalk-like viruses in Germany.";

RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF427118; AAL18874.1; -

DR InterPro; IPR004005; Calici\_coat.

DR InterPro; IPR008975; Viral\_cap\_coat.

DR Pfam; PF00915; Calici\_coat; 1.

SQ SEQUENCE 548 AA; 59874 MW; 9A1D8BEE58ADE7D CRC64;

Query Match 70.4%; Score 2016; DB 12; Length 548;

Best Local Similarity 68.9%; Pred. No. 3.7e-150;

Matches 381; Conservative 73; Mismatches 83; Indels 16; Gaps 8;

QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLPVPVAGASIAAPVAGQNIIDPWIRTFVQAP 60

DB 1 MKMASNDAAPSDCAAGLVPEESNVEVMALEPVVAGASIAAPVGTQNIIDPWIRMFVQAP 60

QY 61 NGFTVSPRNSPGEIILNLELGPDLNPLYLAHLSRMVNGYAGGVQVLLAGNAFTAGKIL 120

DB 61 NGFTVSPRNSPGEVLLNLELGPDLNPLYLAHLSRMVNGYAGGVQVLLAGNAFTAGKLI 120

QY 121 FAATPPNPLVDMLSPAQITMLPHLIVDRTLEPMTPLDVRNVFTHFNQPPRMRLVA 180

DB 121 FAATPPNPLVDMLSPAQITMLPHLIVDRTLEPMTPLDVRNVFTHFNQPPRMRLVA 180

QY 181 MLYTFLRNSGDDVFTVSCRVLTPTDPDEFIYLVPPSVESKTPFTPLITISLTNS 240

DB 181 MLYTFLRNSGDDVFTVSCRVLTPTDPDEFIYLVPPSVESKTPFTPLITISLTNS 240

QY 241 RPIPIEOLYAPNETNVVQNGRCCTLDGELQGTQTLSSAVCFLOGRT---VAD--NG 295

DB 241 RPIPIEOLYAPNETNVVQNGRCCTLDGELQGTQTLSSAVCFLOGRT---VAD--NG 295

QY 296 DNWDQNLQLYTPNGASVDPDTPVAPLCTODFGMLYGVLTQDNVNVSTGE---AKNA 351

DB 296 DNWDQNLQLYTPNGASVDPDTPVAPLCTODFGMLYGVLTQDNVNVSTGE---AKNA 351

QY 301 THWNVTNVL---NGTDPDPTEDVPAPLGTDFSGIYGVISQRNTVTPGEGDLPANRA 357

DB 301 THWNVTNVL---NGTDPDPTEDVPAPLGTDFSGIYGVISQRNTVTPGEGDLPANRA 357

QY 352 KGIYSTSGKTPKIGSIGLSH-ITEVHPNQOSRFTPVGVA-VDENTPFOQWLPHYA 409

DB 352 KGIYSTSGKTPKIGSIGLSH-ITEVHPNQOSRFTPVGVA-VDENTPFOQWLPHYA 409

QY 358 HEAVIATYSPKTPKLGNIQFSTWETQDVSSQPTKFTPVGLASVDANSFHDQWTLPSYS 417

DB 358 HEAVIATYSPKTPKLGNIQFSTWETQDVSSQPTKFTPVGLASVDANSFHDQWTLPSYS 417

QY 410 GSLALNTNLAPAVATPFGQQLLFRSRVPCVQLOGODAFIDCLLPQEWNVHYQEAAP 469

DB 410 GSLALNTNLAPAVATPFGQQLLFRSRVPCVQLOGODAFIDCLLPQEWNVHYQEAAP 469

QY 418 GALTLMNLNAPSAPVFPCECLLFRSFIP-LKGGYGNA-IDCLMPQEWQHLQESAP 475

DB 418 GALTLMNLNAPSAPVFPCECLLFRSFIP-LKGGYGNA-IDCLMPQEWQHLQESAP 475

QY 470 SQADVALIRVNPDTGRTLFEAKLHRSRGTIVSHTGAYPLVVPNGHFRFDSWVNYSL 529

DB 470 SQADVALIRVNPDTGRTLFEAKLHRSRGTIVSHTGAYPLVVPNGHFRFDSWVNYSL 529

QY 476 SLSDVALIRVNPETGRTLFEAKLHRSRGTIVSHTGAYPLVVPNGHFRFDSWVNYSL 535

DB 476 SLSDVALIRVNPETGRTLFEAKLHRSRGTIVSHTGAYPLVVPNGHFRFDSWVNYSL 535

QY 530 APMTGNGRRRIQ 542

DB 530 APMTGNGRRRIQ 542

QY 536 APWNGSGRRRMQ 548

DB 536 APWNGSGRRRMQ 548

RESULT 10

Q917Y9 PRELIMINARY; PRT; 535 AA.

AC Q917Y9; (TrEMBLrel. 19, Created)

DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)

DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)

DE Capsid protein.

OS Norwalk-like virus NLV/Miami/81/1986/US.

OC Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;

OC Norovirus.

OC NCBI\_TaxID=171848;

OX [1]

SEQUENCE FROM N.A.

RC STRAIN=Hu/NLV/Miami/81/1986/US;

RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;

RT "A one-tube method of reverse transcription-PCR to efficiently amplify a 3-kilobase region from the RNA polymerase gene to the poly(A) tail of small round-structured viruses (Norwalk-like viruses).";

RL J. Clin. Microbiol. 35:570-577(1997).

RN [2]

SEQUENCE FROM N.A.

RC STRAIN=Hu/NLV/Miami/81/1986/US;

RA Ando T., Noel J.S., Fankhauser R.L.;

RT "Genetic classification of 'Norwalk-like viruses.'";

RL J. Infect. Dis. 181:S336-S348(2000).

RN [3]

SEQUENCE FROM N.A.

RC STRAIN=Hu/NLV/Miami/81/1986/US;

RA Ando T., Seto Y., Noel J.S., Monroe S.S., Fankhauser R.L.;

RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF414416; AAL13001.1; -

DR InterPro; IPR004005; Calici\_coat.

DR InterPro; IPR008975; Viral\_cap\_coat.

DR Pfam; PF00915; Calici\_coat; 1.

SQ SEQUENCE 535 AA; 56745 MW; D43CDF7BD8C7225A CRC64;

Query Match 70.4%; Score 2015.5; DB 12; Length 535;

Best Local Similarity 71.8%; Pred. No. 3.9e-150;

Matches 390; Conservative 61; Mismatches 83; Indels 9; Gaps 6;

QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLPVPVAGASIAAPVAGQNIIDPWIRTFVQAP 60

DB 1 MKMASNDAAPSNDGAAGLVPEVNETMALEPVVAGASIAAPLTGQNNVIDPWIRMFVQAP 60

QY 61 NGFTVSPRNSPGEIILNLELGPDLNPLYLAHLSRMVNGYAGGVQVLLAGNAFTAGKIL 120

DB 61 NGFTVSPRNSPGEIILNLELGPDLNPLYLAHLSRMVNGYAGGVQVLLAGNAFTAGKLV 120

QY 121 FAATPPNPLVDMLSPAQITMLPHLIVDRTLEPMTPLDVRNVFTHFNQPPRMRLVA 180

DB 121 FAATPPNPLVDMLSPAQITMLPHLIVDRTLEPMTPLDVRNVFTHFNQPPRMRLVA 180

QY 181 MLYTFLRNSGDDVFTVSCRVLTPTDPDEFIYLVPPSVESKTPFTPLITISLTNS 240

DB 181 MLYTFLRNSGDDVFTVSCRVLTPTDPDEFIYLVPPSVESKTPFTPLITISLTNS 240

QY 241 RPIPIEOLYAPNETNVVQNGRCCTLDGELQGTQTLSSAVCFLOGRTVADNNGDNWDQ 300

DB 241 RPIPIEOLYAPNETNVVQNGRCCTLDGELQGTQTLSSAVCFLOGRTVADNNGDNWDQ 300

QY 241 RFPVPIDELATYSPNEGLVQVQNGRSTLDGELGTTQLVPSNICSLRGRINAHLSNQHR 300

DB 241 RFPVPIDELATYSPNEGLVQVQNGRSTLDGELGTTQLVPSNICSLRGRINAHLSNQHR 300

QY 301 NLLQLYTPNGASVDPDTPVAPLGTDFSGMLYGVLTQDNVNVSTGEAKNAGIYISTTS 360

DB 301 NLLQLYTPNGASVDPDTPVAPLGTDFSGMLYGVLTQDNVNVSTGEAKNAGIYISTTS 360

QY 301 WNVQTNANGTTPDPTEDVPAPLGTDFLANIYGVTSQRNPD-NTCRAHD--GI-LATWS 356

DB 301 WNVQTNANGTTPDPTEDVPAPLGTDFLANIYGVTSQRNPD-NTCRAHD--GI-LATWS 356

QY 361 GKFTPKIGSIGLHSEITEH-VHPNQOSRFTPVGVAVDENTPFOQWLPVHAGSLALNTNLA 419

DB 361 GKFTPKIGSIGLHSEITEH-VHPNQOSRFTPVGVAVDENTPFOQWLPVHAGSLALNTNLA 419

QY 357 PKFTPKLGSVVLGTWEDRDFDQNPTRFTPTPVGLY--DTDFNQWLPVYSGALTLNNLA 414

DB 357 PKFTPKLGSVVLGTWEDRDFDQNPTRFTPTPVGLY--DTDFNQWLPVYSGALTLNNLA 414

```
QY 420 PAVAPTFEGEQLLFFRSRVPVCGVQGGQDAFIDCILLPOEWNHFFYQEAAPSOADVALIRY 479
Db 415 PSVAPLFFGEEQLLFFRSRVPVCGGT--SNGAIDCILLPOEWNHFFYQESAPSSTVDALIRY 472
QY 480 VNPDGRTLFEAKLHRSRGPITVSHGTGAYPLVPPNGHFRFDSWVWVNFYSLAPMGNGRR 539
Db 473 TNPDGRTLFEAKLHRSRGPITVSHGTGAYPLVPPNGHFRFDSWVWVNFYSLAPMGNGRR 532
QY 540 RIQ 542
Db 533 RVQ 535

RESULT 11
Q917Y3
ID Q917Y3 PRELIMINARY; PRT; 535 AA.
AC Q917Y3;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Westover/302/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171850;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Westover/302/1994/US;
RX MEDLINE=97193806; PubMed=9401391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Westover/302/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
characterized small round-structured viruses involved in outbreaks of
nonbacterial acute gastroenteritis in the United States, 1990 to
1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Westover/302/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Westover/302/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL, AF414418, RAL13007.1;
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 535 AA; 58684 MW; 62620468ABB44EC0 CRC64;

Query Match 70.2%; Score 2010.5; DB 12; Length 535;
Best Local Similarity 71.6%; Pred. No. 9.6e-150;
Matches 389; Conservative 61; Mismatches 84; Indels 9; Gaps 6;

QY 1 MMSANDAPSDNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNTIIDPWRITNFVQAP 60
Db 1 MMSANDAPSDNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNTIIDPWRITNFVQAP 60
QY 61 NGEFTVSPRNSPGEILLNLELPGDLPNLYLAHLRMYNGYAGVGVQVLLAGNAFTAGKIL 120
```

```
Db 61 NGEFTVSPRNSPGEILLNLELPGDLPNLYLAHLRMYNGYAGVGVQVLLAGNAFTAGKIL 120
QY 121 FAALPPNPLVDMISPAQITMLPHLIVDVRTLEPIMTLPDVRNVFFYNNQOPQMRMLVA 180
Db 121 FAALPPHPIENLSPGQITMFPVHVIDVRLPVLPLPDVRNFFHYNQBPPEPRMLVA 180
QY 181 MLYPLRNSGGDDVFTVSCRVLTPTPTDFEFYLYVPPSVESKTPFTLPILITISLNTS 240
Db 181 MLYPLRNSGGDDVFTVSCRVLTPTPTDFEFYLYVPPSVESKTPFTLPILITISLNTS 240
QY 241 RPIPIEOLYAPNETNVCQNGRCITLDGELQGTQTLSSAVCFLOGRTVADNGDNMDO 300
Db 241 RFPAPIDELYTPNEGLVVPQNGRSTLDGELLGTTLQVPSNICSLRGRIHAHLSDNOHR 300
QY 301 NLLQLYTPNGASYDPTDEVPAPLGTQDFSGMLYGLVLTQDNVNVSTGEAKNAKGIYISITS 360
Db 301 WNMQVTNANGTPDPTEDVPAPLGTDFLANIYGVTSQRNPD-NTCRAHD--GI-LATWS 356
QY 361 GKFTPKIGSIGLHSITEH-VHPNQSRPTPGVAVDENTPFOQWVLPVHYAGSLANTNLA 419
Db 357 PKFTPKLGSVVLGTWEDRDFDINOPTPTPVGLY--DTDHFNQWALPNYSGALTLMNLA 414
QY 420 PAVAPTFEGEQLLFFRSRVPVCGVQGGQDAFIDCILLPOEWNHFFYQEAAPSOADVALIRY 479
Db 415 PSVAPLFFGEEQLLFFRSRVPVCGGT--SNGAIDCILLPOEWNHFFYQESAPSSTVDALIRY 472
QY 480 VNPDGRTLFEAKLHRSRGPITVSHGTGAYPLVPPNGHFRFDSWVWVNFYSLAPMGNGRR 539
Db 473 TNPDGRTLFEAKLHRSRGPITVSHGTGAYPLVPPNGHFRFDSWVWVNFYSLAPMGNGRR 532
QY 540 RIQ 542
Db 533 RVQ 535

RESULT 12
Q917X4
ID Q917X4 PRELIMINARY; PRT; 535 AA.
AC Q917X4;
DT 01-DEC-2001 (TREMELrel. 19, Created)
DT 01-DEC-2001 (TREMELrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Port Canaveral/301/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171853;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;
RX MEDLINE=97193806; PubMed=9401391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
characterized small round-structured viruses involved in outbreaks of
nonbacterial acute gastroenteritis in the United States, 1990 to
1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Port Canaveral/301/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
```

Ando T., Noel J.S., Fankhauser R.L.;
"Genetic classification of Norwalk-like viruses.";
J. Infect. Dis. 181:S336-S348(2000).
[4]
SEQUENCE FROM N.A.
STRAIN=Hu/NLW/Port Canaveral/301/1994/US;
Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
EMBL; AF414421; AAL13016.1; -
InterPro; IPR004005; Calici.coat.
InterPro; IPR008975; Viral Cap.coat.
Pfam; PF00915; Calici.coat.1;
SEQUENCE 535 AA; 58738 MW; 7C29B02DFDC724C6 CRC64;
Query Match 70.1%; Score 2007.5; DB 12; Length 535;
Best Local Similarity 71.5%; Pred. No. 1.7e-149;
Matches 388; Conservative 62; Mismatches 84; Indels 9; Gaps 535
QY 1 MKMASNDAAPSNDGAASIVPEGINETMPLEPVAGASIAAPVAGQTNIIDPMTRTNFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEVNNETMALPEPVAGASIAAPLGQNVIDPWRMNFVQAP 60
QY 61 NGEFVTSPRNSPGGILLNLNLELGPDLNPYLHLHSLRMNYNGVAGGVVEOVLLAGNAFTAGKIL 120
Db 61 NGEFVTSPRNSPGGILLNLNLELGPDLNPFLHLHSLRMNYNGVAGGVVEOVLLAGNAFTAGKL 120
QY 121 FAATPPNFLVDMI SPAQITMLPHLIVDRTLEIPMTPLPDVRNVVFHNNOQPQRMLVA 180
Db 121 FAATPPHPFIENLSFGQITMPFHVIIDVRTLEPLLPLPDVRNNFFHYNQEPERMLVA 180
QY 181 MLYTPLRNSGGDDVFTVSCRVLTRPTDPDEFVLVPVSVESKTKEFTLPILTISLTNS 240
Db 181 MLYTPLRNSGGDDVFTVSCRVLTRPSDFDFNLVPPTVESKTKEFTLPILTIGELNS 240
QY 241 RPPIEQLIYTAPNETNIVQONGRCITLDGELOTTTOLLSSAVCFLOGRTVADGNWDQ 300
Db 241 RPFAIDELYTSNGLVVPQNGCRPTLDGELLGTQLVPSNTCSLRGRINAHLSDNQHR 300
QY 301 NULLTYPNGASYDTDEVPAPLGTQDSFMLXGVLTDQNNVNSTGEAKNAGIYSTS 360
Db 301 WNMQVTNANGTPFDPTEDAPAPLGTDFLANIYGVTYSQRMPD-NTRCHD--GI-LATWS 356
QY 361 GKFTPKIGSIGLHSITEH-VHPNQOSRPTPGVAVDENTPFQQWLPHYAGSLANTNLA 419
Db 357 PKFTPKIGSVLCTWEDRDFDIQNRTRTPVGLY--DTDHENQWALPNYSALTANMLA 414
QY 420 PAVAPTFPGEQLLFRRSPVCVQGSGODAFIDCLLPOEWNNHFYQBAASQADVAILRY 479
Db 415 PSVAPLFPGEQLLFRRSHIPLKGGT--SNGAIDCLLPOEWYQHFYQSAPSSTDVALIRY 472
QY 480 VNPDTGRTLFEAKLHRSRGITVSHTGAYPLVVYPNGHFRPDSWNQFYSIAPMGTCNGRR 539
Db 473 TNPDTRGLVFEAKLHQRFQFITVANSGSRPIVVPPNGYFRFDSWNQFYSIAPMGTCNGRR 532
QY 540 RIQ 542
Db 533 RVQ 535
RESULT 13
Q916E4 PRELIMINARY; PRT; 535 AA.
ID Q916E4
AC Q916E4; (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/Wiesbaden 294/01/DE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Calciviridae;
OC Norovirus.
OX NCBI_TaxID=173922;
RP SEQUENCE FROM N.A.

RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
RT of small round-structured viruses (Norwalk-like viruses).";  
RL J. Clin. Microbiol. 35:570-577(1997).  
RN [2]  
RP SEQUENCE OF 100-192 FROM N.A.  
RC STRAIN=Hu/NLV/Richmond/283/1994/US;  
RX MEDLINE=98071277; PubMed=9407386;  
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,  
RA Seto Y., Monroe S.S., Glass R.I.;  
RT "Correlation of patient immune responses with genetically  
RT characterized small round-structured viruses involved in outbreaks of  
RT nonbacterial acute gastroenteritis in the United States, 1990 to  
RT 1995.";  
RL J. Med. Virol. 53:372-383(1997).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Richmond/283/1994/US;  
RX MEDLINE=20266071; PubMed=10804147;  
RA Ando T., Noel J.S., Fankhauser R.L.;  
RT "Genetic classification of 'Norwalk'-like viruses.";  
RL J. Infect. Dis. 181:S336-S348(2000).  
RN [4]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Richmond/283/1994/US;  
RA Ando T., Seto Y., Noel J.S., Glass R.I., Fankhauser R.L.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF414419; AAL13010.1; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 535 AA; 58742 MW; 4EDACBCEB22F80 CRC64;

Query Match 70.0%; Score 2003.5; DB 12; Length 535;  
Best Local Similarity 71.5%; Pred. No. 3.4e-149; Indels 9; Gaps 6;  
Matches 388; Conservative 62; Mismatches 84;  
QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNTIIDPWIRTFVQAP 60  
DB 1 MKMASNDAAPSNDGAAGLVPEVNNETMALEPVAGASIAAPLTGQNNVIDPWIRMFVQAP 60  
QY 61 NGEFTVSPRNSPGEILLNLLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120  
DB 61 NGEFTVSPRNSPGEILLNLLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120  
QY 121 FAALPPNLFVDMSPAQITMLPHLIVDVRTLEPIMTLPDVRNVFHHNQOPQPMRLVA 180  
DB 121 FAALPPHPIENLSPGQITMFPVHVIDVRTLEPVLPLPVDVRNFFHYNQPEPRMRLVA 180  
QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEIYLVPPSVESKTKPFTLPILTIGELTNS 240  
DB 181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEIYLVPPSVESKTKPFTLPILTIGELTNS 240  
QY 241 RFPPIEQLYTAPNETNVVQCGRCGLDGELOQTQLLSSAVCFLOQRTVADNGDNWDQ 300  
DB 241 RFPAPIDELYTSNPEGLVVPQNGSTLDGELLGTTQLVPSNICSLRGRINAHLSDNQHR 300  
QY 301 NLLQLYTPNGASYDPTDEVPAPLGTDQFSGLMYGVLTQDNVNVSTGEAKNAGIYISTTS 360  
DB 241 RFPAPIDELYTSNPEGLVVPQNGSTLDGELLGTTQLVPSNICSLRGRINAHLSDNQHR 300  
QY 301 NLLQLYTPNGASYDPTDEVPAPLGTDQFSGLMYGVLTQDNVNVSTGEAKNAGIYISTTS 360  
DB 301 NLLQLYTPNGASYDPTDEVPAPLGTDQFSGLMYGVLTQDNVNVSTGEAKNAGIYISTTS 360  
QY 301 NLLQLYTPNGASYDPTDEVPAPLGTDQFSGLMYGVLTQDNVNVSTGEAKNAGIYISTTS 360  
DB 301 NLLQLYTPNGASYDPTDEVPAPLGTDQFSGLMYGVLTQDNVNVSTGEAKNAGIYISTTS 360  
QY 361 GKFTPKIGSLGHSITEH-VHPNQSRPTPGVAVDENTPFOQWVLPYVAGSLALNTMLA 419  
DB 361 GKFTPKIGSLGHSITEH-VHPNQSRPTPGVAVDENTPFOQWVLPYVAGSLALNTMLA 419  
QY 419 PSVAPLPFGEQLLFFRSHRSHPLKGGT--SNGAIDCLLPQEWVQHVFQESAPSTVDALLRY 472  
DB 419 PSVAPLPFGEQLLFFRSHRSHPLKGGT--SNGAIDCLLPQEWVQHVFQESAPSTVDALLRY 472  
QY 480 VNPDGRTLFEAKLHRSRSGFITVSHTGAYPLVVPNGHFRFDSWVNVQFYSLAPMGNGRR 539  
DB 480 VNPDGRTLFEAKLHRSRSGFITVSHTGAYPLVVPNGHFRFDSWVNVQFYSLAPMGNGRR 539  
QY 473 TNPDTGRVLFEAKLHRSRSGFITVSHTGAYPLVVPNGHFRFDSWVNVQFYSLAPMGNGRR 532  
DB 473 TNPDTGRVLFEAKLHRSRSGFITVSHTGAYPLVVPNGHFRFDSWVNVQFYSLAPMGNGRR 532

QY 540 RIQ 542  
DB 533 RVQ 535  
RESULT 15  
Q916E6 PRELIMINARY; PRT; 535 AA.  
AC Q916E6;  
DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus NLV/Dillingen 391/01/DE.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=173920;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NLV/Dillingen 391/01/DE;  
RA Kuenkel U., Hoehne M., Schreier E.;  
RT "Molecular epidemiology of outbreaks of gastroenteritis associated  
RT with Norwalk-like viruses in Germany.";  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF425767; AAL18860.1; -;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 535 AA; 58685 MW; 0520E521A786734A CRC64;

Query Match 69.9%; Score 2002.5; DB 12; Length 535;  
Best Local Similarity 71.3%; Pred. No. 4.1e-149;  
Matches 387; Conservative 62; Mismatches 85; Indels 9; Gaps 6;  
QY 1 MKMASNDAAPSNDGAASLVPEGINETMPLEPVAGASIAAPVAGQNTIIDPWIRTFVQAP 60  
DB 1 MKMASNDAAPSNDGAAGLVPEVNNETMALEPVAGASIAAPLTGQNNVIDPWIRMFVQAP 60  
QY 61 NGEFTVSPRNSPGEILLNLLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120  
DB 61 NGEFTVSPRNSPGEILLNLLELGPDLNPLYLAHLSRMVYAGVGVQVLLAGNAFTAGKIL 120  
QY 121 FAALPPNLFVDMSPAQITMLPHLIVDVRTLEPIMTLPDVRNVFHHNQOPQPMRLVA 180  
DB 121 FAALPPHPIENLSPGQITMFPVHVIDVRTLEPVLPLPVDVRNFFHYNQPEPRMRLVA 180  
QY 181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEIYLVPPSVESKTKPFTLPILTIGELTNS 240  
DB 181 MLYTPLRSNGSGDDVFTVSCRVLTRPTDPFEIYLVPPSVESKTKPFTLPILTIGELTNS 240  
QY 241 RFPPIEQLYTAPNETNVVQCGRCGLDGELOQTQLLSSAVCFLOQRTVADNGDNWDQ 300  
DB 241 RFPAPIDELYTSNPEGLVVPQNGSTLDGELLGTTQLVPSNICSLRGRINAHLSDNQHR 300  
QY 301 NLLQLYTPNGASYDPTDEVPAPLGTDQFSGLMYGVLTQDNVNVSTGEAKNAGIYISTTS 360  
DB 301 NLLQLYTPNGASYDPTDEVPAPLGTDQFSGLMYGVLTQDNVNVSTGEAKNAGIYISTTS 360  
QY 301 NLLQLYTPNGASYDPTDEVPAPLGTDQFSGLMYGVLTQDNVNVSTGEAKNAGIYISTTS 360  
DB 301 NLLQLYTPNGASYDPTDEVPAPLGTDQFSGLMYGVLTQDNVNVSTGEAKNAGIYISTTS 360  
QY 361 GKFTPKIGSLGHSITEH-VHPNQSRPTPGVAVDENTPFOQWVLPYVAGSLALNTMLA 419  
DB 361 GKFTPKIGSLGHSITEH-VHPNQSRPTPGVAVDENTPFOQWVLPYVAGSLALNTMLA 419  
QY 419 PSVAPLPFGEQLLFFRSHRSHPLKGGT--SNGAIDCLLPQEWVQHVFQESAPSTVDALLRY 472  
DB 419 PSVAPLPFGEQLLFFRSHRSHPLKGGT--SNGAIDCLLPQEWVQHVFQESAPSTVDALLRY 472  
QY 480 VNPDGRTLFEAKLHRSRSGFITVSHTGAYPLVVPNGHFRFDSWVNVQFYSLAPMGNGRR 539  
DB 480 VNPDGRTLFEAKLHRSRSGFITVSHTGAYPLVVPNGHFRFDSWVNVQFYSLAPMGNGRR 539  
QY 473 TNPDTGRVLFEAKLHRSRSGFITVSHTGAYPLVVPNGHFRFDSWVNVQFYSLAPMGNGRR 532  
DB 473 TNPDTGRVLFEAKLHRSRSGFITVSHTGAYPLVVPNGHFRFDSWVNVQFYSLAPMGNGRR 532  
QY 540 RIQ 542  
DB 533 RVQ 535

Wed Jun 2 09:13:35 2004

Db 533 RVQ 535

Search completed: June 1, 2004, 13:53:21  
Job time : 32.556 secs

us-09-926-799-9.rsp

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 47.2483 Seconds  
(without alignments)

3289.030 Million cell updates/sec

Title: US-09-926-799-10

Perfect score: 2906

Sequence: 1 MKMASNDAPNSDGAANLVP.....GNQPYTLAPMGSGGRRRAQ 550

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04.\*  
1: Geneseqp1980s.\*  
2: Geneseqp1990s.\*  
3: Geneseqp2000s.\*  
4: Geneseqp2001s.\*  
5: Geneseqp2002s.\*  
6: Geneseqp2003as.\*  
7: Geneseqp2003bs.\*  
8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2906	100.0	550	4	AAB49709 Small rou
2	2124	73.1	548	4	AAB49705 Small rou
3	2074	71.4	548	5	AAU91272 Norwalk v
4	2062	71.0	535	4	AAB49707 Small rou
5	2001	68.9	540	4	AAB49706 Small rou
6	1872	64.4	542	4	AAB49708 Small rou
7	1834.5	63.1	539	4	AAB49704 Small rou
8	1613.5	55.5	541	4	AAB49710 Small rou
9	1154.5	39.7	545	4	AAB49700 Small rou
10	1141	39.3	530	2	AAR50972 Norwalk v
11	1141	39.3	530	7	ADC72176 Norwalk v
12	1131	38.9	530	4	AAB49701 Small rou
13	1124.5	38.7	544	4	AAB49703 Small rou
14	1097	37.7	530	2	AAR57091 Small rou
15	1090.5	37.5	546	4	AAB49702 Small rou
16	285	9.8	579	2	AAW08143 RHDV caps
17	280.5	9.7	547	4	AAM50108 Feline ca
18	280.5	9.7	668	4	AAB67462 Amino aci
19	280.5	9.7	671	4	AAM50107 Feline ca
20	276.5	9.5	622	4	AAB47045 Feline Ca
21	276.5	9.5	623	4	AAB47044 Feline Ca
22	276.5	9.5	623	4	AAB47043 Feline Ca
23	272.5	9.4	669	4	AAB67461 Amino aci
24	270.5	9.3	668	2	AAR10686 Feline ca
25	270.5	9.3	668	4	AAE04304 Feline ca

26	179.5	6.2	40	5	AAU91274	AAU91274 Norwalk v
27	168	5.8	40	5	AAU91273	AAU91273 Norwalk v
28	138.5	4.4	934	1	AAP20016	AAP20016 Sequence
29	124.5	4.3	1147	5	ABB76724	ABB76724 Foot and
30	121	4.2	1344	4	RAG93134	RAG93134 C glutami
31	120	4.1	40	5	AAU91275	AAU91275 Norwalk v
32	119.5	4.1	2608	5	ABG70209	ABG70209 Human pre
33	119	4.1	1036	6	ABU70541	ABU70541 Human adi
34	119	4.1	2647	2	AAW19349	AAW19349 Human fil
35	119	4.1	2647	2	AAV33867	AAV33867 Carboxyl
36	119	4.1	2647	6	ABR59725	ABR59725 Human fil
37	117	4.0	4618	4	AAW39043	AAW39043 Human pol
38	113	3.9	701	4	ABB57994	ABB57994 Drosophil
39	112.5	3.9	568	4	ABB64772	ABB64772 Drosophil
40	109.5	3.8	1079	2	AAV03163	AAV03163 MuLV reve
41	109.5	3.8	1199	3	AAW12994	AAW12994 MLV rever
42	109.5	3.8	1224	2	AAV17947	AAV17947 MoMLV pol
43	109.5	3.8	1737	3	AAW10044	AAW10044 MMLV 989-
44	107.5	3.7	1958	4	ABG21921	ABG21921 Novel hum
45	107	3.7	557	5	ABP61060	ABP61060 Lactobaci

#### ALIGNMENTS

RESULT 1  
AAB49709  
ID AAB49709 standard; protein; 550 AA.  
XX  
AC AAB49709;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 10.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
FD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX  
PA (NTNA-) JAPAN NAT INST INFECTIOUS DISEASES.  
FA (DENK-) DENKA SEIKEN KK.  
XX  
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
WPI; 2001-080848/09.  
DR N-PSDB; AAF29150.  
XX  
Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.  
XX  
Claim 1; Page 62-64; 84pp; Japanese.  
XX  
This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

Sequence 550 AA;

Query Match 100.0%; Score 2906; DB 4; Length 550;  
Best Local Similarity 100.0%; Pred. No. 1.4e-260;

Matches	550;	Conservative	0;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1	MMKASNDAA	PSNDGAANL	VPPEANDE	VMALBPVVGASIA	APVVGQNI	IDPWIRENF	VQAP	60
Db	1	MMKASNDAA	PSNDGAANL	VPPEANDE	VMALBPVVGASIA	APVVGQNI	IDPWIRENF	VQAP	60
QY	61	QGEFTVSP	RNSPGEM	LNLELGP	ELNLPVLSHLSRM	YAGGMQVQVVL	AGNAFTAGKII	120	
Db	61	QGEFTVSP	RNSPGEM	LNLELGP	ELNLPVLSHLSRM	YAGGMQVQVVL	AGNAFTAGKII	120	
QY	121	FAAVPPHP	PPVENISAAQITMC	PHIVDVRQLEP	VLLPLPDIRNRP	FFHYNQENTPRM	ELVA	180	
Db	121	FAAVPPHP	PPVENISAAQITMC	PHIVDVRQLEP	VLLPLPDIRNRP	FFHYNQENTPRM	ELVA	180	
QY	181	MLYTPLR	ANSGBD	VFTVSCR	VLTRPAPDFE	FTFLVPVTSKTP	FTFLPILITGELS	NSR	240
Db	181	MLYTPLR	ANSGBD	VFTVSCR	VLTRPAPDFE	FTFLVPVTSKTP	FTFLPILITGELS	NSR	240
QY	241	FPAADML	YTDNPNES	IVVQPNQGR	CTLDGTTLQ	GTTLQVPTQICAPRGT	LISQTA	RAADST	300
Db	241	FPAADML	YTDNPNES	IVVQPNQGR	CTLDGTTLQ	GTTLQVPTQICAPRGT	LISQTA	RAADST	300
QY	301	DSFQAR	NHPLHVQVKNL	DGTYDPTDDIP	AVLGAIDFKG	TVFGVASQ	RDVSGQ	EQEGHY	360
Db	301	DSFQAR	NHPLHVQVKNL	DGTYDPTDDIP	AVLGAIDFKG	TVFGVASQ	RDVSGQ	EQEGHY	360
QY	361	ATRAHEA	HDITDTPK	YAPKGLTILIKSGS	DDNTNQPIRFT	TPVGMG	DDNNRQWEL	PDYSG	420
Db	361	ATRAHEA	HDITDTPK	YAPKGLTILIKSGS	DDNTNQPIRFT	TPVGMG	DDNNRQWEL	PDYSG	420
QY	421	RLTLNN	MLAPAVSPSPG	ERILFRSIV	PSAGYSGYID	CLIPQEWQHF	YQEAAPSQS	480	
Db	421	RLTLNN	MLAPAVSPSPG	ERILFRSIV	PSAGYSGYID	CLIPQEWQHF	YQEAAPSQS	480	
QY	481	AVALVRY	NPD	TGRNIFEAKLH	REGFLTVANC	GNPIVVP	PNGYFRFEAWGNQFY	TLAPM	540
Db	481	AVALVRY	NPD	TGRNIFEAKLH	REGFLTVANC	GNPIVVP	PNGYFRFEAWGNQFY	TLAPM	540
QY	541	GSQGR	RRRAQ	550					
Db	541	GSQGR	RRRAQ	550					
RESULT 2									
ID	AAAB49705								
XX	AAAB49705	standard; protein; 548 AA.							
AC	AAAB49705;								
DT	04-APR-2001	(first entry)							
DE	Small round structured virus protein	SEQ ID 6.							
XX	Small round structured virus; SRSV; food poisoning.								
XX	Small round structured virus.								
OS	WO200079280-A1.								
PN	28-DEC-2000.								
XX	22-JUN-2000;	2000WO-JP004095.							
PR	22-JUN-1999;	99JP-00175928.							
XX	(NINA-) JAPAN NAT INST INFECTIOUS DISEASES.								
PA	(DENK-) DENKA SEIKEN KK.								
XX	Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;								
XX	WPI; 2001-080848/09.								
DR	N-PSDB; AAF29146.								
XX									

PT	Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.											
PT												
PT												
XX	Claim 1; Page 52-54; 84pp; Japanese.											
PS												
XX	This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks											
CC												
CC												
XX	Sequence 548 AA;											
SQ												
QY	Query Match 73.1%; Score 2124; DB 4; Length 548;											
Db	Best Local Similarity 71.4%; Pred. No. 6.3e-188;											
QY	Matches 396; Conservative 72; Mismatches 75; Indels 12; Gaps 4											
Db												
QY	1 MMKASNDAA	PSNDGAANL	VPPEANDE	VMALBPVVGASIA	APVVGQNI	IDPWIRENF	VQAP	60				
Db	1 MMKASNDAA	PSNDGAAGL	VPPEINNEA	MDPVAGAA	LAAPLT	QGOQNI	IDPWIMNN	FVQAP	60			
QY	61 QGEFTVSP	RNSPGEM	LNLELGP	ELNLPVLSHLSRM	YAGGMQVQVVL	AGNAFTAGKII	120					
Db	61 GGEFTVSP	RNSPGEV	LNLELGP	ELNLPYLAH	ARMYNGYAGG	FEVQVLAGNAFTAGKII	120					
QY	121 FAAVPPHP	PPVENISAAQITMC	PHIVDVRQLEP	VLLPLPDIRNRP	FFHYNQENTPRM	ELVA	180					
Db	121 FAALPPNP	PPIDNLSAAQITMC	PHIVDVRQLEP	VLLPMPDVRN	FFHYNQNGSD	RLRLTA	180					
QY	181 MLYTPLRA	NSGBD	VFTVSCR	VLTRPAPDFE	FTFLVPVTSKTP	FTFLPILITGELS	NSR	239				
Db	181 MLYTPLRANS	GDDVFTVSCR	VLTRPSPDF	SFNFVPTVESK	TKDFTLPI	LITISEMS	240					
QY	240 RFPAAID	MLYTDN	PNESIVVQPNQGR	CTLDGTTLQ	GTTLQVPTQICAPRGT	LISQTA	RAADS	299				
Db	241 RFPVP	IESLHTSP	ENIVVQCNQGR	VTLDELGM	GTTLQPSQICAPRGT	LRSTRS	RAADQ	300				
QY	300 TD-SFQAR	NHPLHVQVKNL	DGTYDPTDDIP	AVLGAIDPKG	TVFGVASORDV	SGOQEG	358					
Db	301 ADTAT	PRLFN	VYWHVQ	LDNLNGT	PDPAEDIPG	PLGTPDPRG	KVFGVASQR	NLDS	355			
QY	359 HYATRA	HEAHDITD	TPKAPK	GLTILIKSGS	DDNTNQPIRFT	TPVGMG	DDNN---	WRQWEL	415			
Db	356 --TT	RAHEAKVDIT	TAGRTPK	LSLEISTDS	DDQDQNTKFT	FPVGLGV	DN	EA	FQWLSL	413		
QY	416 PDYSG	RLTLNN	MLAPAVSPSPG	ERILFRSIV	PSAGYSGYID	CLIPQEWQHF	YQEA	475				
Db	414 PDYSG	QFTNN	MLAPAVAPN	FPGEQLJ	FFRSQ	SSGGRNG	VLDCLV	POE	WVQHF	YQES	473	
QY	476 APSQ	SAVALVRY	NPD	TGRNIFEAKLH	REGFLTVANC	GNPIVVP	PNGYFRFEAWGNQFY	530				
Db	474 APAQ	TOVALVRY	NPD	TGKVF	EAKLHKL	GMTI	ANN	GDSP	ITVPPNGYFR	FESVWNP	PFY	530
QY	536 TLAP	WSGGGR	RAQ	550								
Db	534 TLAP	MGTGNGRR	RIQ	548								
RESULT 3												
ID	AAU91272											
XX	AAU91272	standard; protein; 548 AA.										
AC	AAU91272;											
XX	18-JUN-2002	(first entry)										
DE	Norwalk virus associated polynucleotide #1.											
XX	Nowalk virus; monoclonal antibody; geno group I; geno group II;											
KW	immunological detection; food; viral infection.											
KW												

PT	Kit for the detection and typing of small round-structured virus (SRSV)
PT	strains for investigation of food poisoning outbreaks, contains
XX	antibodies.
PS	Claim 1; Page 52-54; 84pp; Japanese.
XX	This invention relates to a kit for the detection and typing of small
CC	round structured virus (SRSV) strains. The kit contains antibodies
CC	directed against peptides represented in sequences AAB49700 - AAB49710,
CC	which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC	AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC	used for detecting and typing strains of SRSV in order to prevent the
CC	spread of infection and to examine the epidemiology of outbreaks
XX	Sequence 548 AA;
SQ	Query Match 73.1%; Score 2124; DB 4; Length 548;
	Best Local Similarity 71.4%; Pred. No. 6.3e-186; Indels 12; Gaps 4;
	Matches 396; Conservative 72; Mismatches 75;
QY	1 MMKASNDAA
Db	1 MMKASNDAA
QY	61 QGEFTVSP
Db	61 QGEFTVSP
QY	121 FAAPVPH
Db	121 FAAPVPH
QY	181 MLYTPLR
Db	181 MLYTPLR
QY	240 RFPAAID
Db	241 RFPVPIE
QY	300 TD-SPO
Db	301 ADTATP
QY	359 HYATRA
Db	356 --TTA
QY	416 PDYSG
Db	414 PDYSG
QY	476 APSQ
Db	474 APAQ
QY	536 TLAP
Db	534 TLAP
RESULT 3	
ID	AAU91272
XX	AAU91272 standard; protein; 548 AA.
AC	AAU91272;
XX	18-JUN-2002 (first entry)
DT	Norwalk virus associated polynucleotide #1.
XX	Nowalk virus; monoclonal antibody; geno group I; geno group II;
KW	immunological detection; food; viral infection.







```
PS Claim 1; Page 59-61; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
CC round structured virus (SRSV) strains. The kit contains antibodies
CC directed against peptides represented in sequences AAB49700 - AAB49710,
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
CC used for detecting and typing strains of SRSV in order to prevent the
CC spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 542 AA;
SQ
Query Match 64.4%; Score 1872; DB 4; Length 542;
Best Local Similarity 64.9%; Pred. No. 1.6e-164;
Matches 361; Conservative 74; Mismatches 101; Indels 20; Gaps 7;
QY 1 MKMASNDAAPSNDGAANLVPEANDVMALEPVVGASIAAPVVGQNIIDPWIRENFVQAP 60
DB 1 MKMASNDAAPSNDGAASLVPGINETMPLEPVVGASIAAPVAGQNIIDPWIRNFVQAP 60
QY 61 QGEFTVSPRNSPGEMLLNLLELGPENLYSLHLSRMVNGYAGGMQVQVVLGNAFTAGKII 120
DB 61 NGEFTVSPRNSPGEMLLNLLELGPENLYSLHLSRMVNGYAGGMQVQVVLGNAFTAGKII 120
QY 121 FAAVPPHPFVENISAAQITMCQPHVIVDVROLEPVLPLDPIRNFPHYQNTPRMRLVA 180
DB 121 FAAVPPHPFVENISAAQITMCQPHVIVDVROLEPVLPLDPIRNFPHYQNTPRMRLVA 180
QY 121 FAAVPPHPFVENISAAQITMCQPHVIVDVROLEPVLPLDPIRNFPHYQNTPRMRLVA 180
DB 121 FAAVPPHPFVENISAAQITMCQPHVIVDVROLEPVLPLDPIRNFPHYQNTPRMRLVA 180
QY 181 MLYTPLRAN-SEGDEVTVCVRLTRPADPDEFTFLVPPTVESKTKPFTLTLTIGELSNS 239
DB 181 MLYTPLRAN-SEGDEVTVCVRLTRPADPDEFTFLVPPTVESKTKPFTLTLTIGELSNS 239
QY 240 RFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLOQTTLQVPTQICAFRGTLSISQARAADS 299
DB 240 RFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLOQTTLQVPTQICAFRGTLSISQARAADS 299
QY 241 RFPPIEQLYTAPNETNVVQCONGRCTLDGELQGTTLQVPTQICAFRGTLSISQARAADS 299
DB 241 RFPPIEQLYTAPNETNVVQCONGRCTLDGELQGTTLQVPTQICAFRGTLSISQARAADS 299
QY 300 TDSQORARNHPLHVQVKNLQDGTQYDPTDDIPAVLGAIDFKGTVFGVASQRDVSGQOQGH 359
DB 300 TDSQORARNHPLHVQVKNLQDGTQYDPTDDIPAVLGAIDFKGTVFGVASQRDVSGQOQGH 359
QY 298 WDQNLQLTYE-----NGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGE-- 347
DB 298 WDQNLQLTYE-----NGASYDPTDEVPAPLGTQDFSGMLYGVLTQDNVNVSTGE-- 347
QY 360 YATRAHEAHDITDPKYAPKLGTLILKSGSDDFNTNQPIRFTPVGMG-DNN--NRWELP 416
DB 360 YATRAHEAHDITDPKYAPKLGTLILKSGSDDFNTNQPIRFTPVGMG-DNN--NRWELP 416
QY 348 -AKNAKGIYITTSKGFPPKIGSLHSITSHVHPNQOSRFTPVGVAVDENTPFQQWVLP 406
DB 348 -AKNAKGIYITTSKGFPPKIGSLHSITSHVHPNQOSRFTPVGVAVDENTPFQQWVLP 406
QY 417 DYSGLTLNMLNLAVALSPSPFGERILFPRSTVPSAGGY--GSGYIDCLIPQEWVQHFQY 474
DB 417 DYSGLTLNMLNLAVALSPSPFGERILFPRSTVPSAGGY--GSGYIDCLIPQEWVQHFQY 474
QY 407 HYAGSLALNTNLAPAVAPTFEGEQLLFFRSRVPVQVQLOGQDAFIDCLLPQEWVQHFQY 466
DB 407 HYAGSLALNTNLAPAVAPTFEGEQLLFFRSRVPVQVQLOGQDAFIDCLLPQEWVQHFQY 466
QY 475 AAPQSQAVLVRYNPDTRGRNIFEAKLHREGFLTVANGCNPNIVVPPNGYFRFPAWGNQF 534
DB 475 AAPQSQAVLVRYNPDTRGRNIFEAKLHREGFLTVANGCNPNIVVPPNGYFRFPAWGNQF 534
QY 467 AAPSQADVALLIRYNPDTRGRNIFEAKLHREGFLTVANGCNPNIVVPPNGYFRFPAWGNQF 526
DB 467 AAPSQADVALLIRYNPDTRGRNIFEAKLHREGFLTVANGCNPNIVVPPNGYFRFPAWGNQF 526
QY 535 YTLAPMGSGQGRRAA 550
DB 535 YTLAPMGSGQGRRAA 550
QY 527 YSLAPMGSGQGRRIQ 542
DB 527 YSLAPMGSGQGRRIQ 542
RESULT 7
AAB49704
ID AAB49704 standard; protein; 539 AA.
XX
XX AAB49704;
AC
XX
XX 04-APR-2001 (first entry)
DT
XX
XX Small round structured virus protein SEQ ID 5.
DE
XX
XX Small round structured virus; SRSV; food poisoning.
KW
XX
XX Small round structured virus.
OS
XX
XX WO200079280-A1.
FN
XX
XX
```

```
PD 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX
XX (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
XX
XX N-FSDB; AAF29145.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX strains for investigation of food poisoning outbreaks, contains
XX antibodies.
XX
XX Claim 1; Page 50-52; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 539 AA;
SQ
Query Match 63.1%; Score 1834.5; DB 4; Length 539;
Best Local Similarity 62.5%; Pred. No. 4.9e-161;
Matches 348; Conservative 85; Mismatches 97; Indels 27; Gaps 7;
QY 1 MKMASNDAAPSNDGAANLVPEANDVMALEPVVGASIAAPVVGQNIIDPWIRENFVQAP 60
DB 1 MKMASNDAAPSNDGAANLVPEANDVMALEPVVGASIAAPVVGQNIIDPWIRENFVQAP 60
QY 61 QGEFTVSPRNSPGEMLLNLLELGPENLYSLHLSRMVNGYAGGMQVQVVLGNAFTAGKII 120
DB 61 QGEFTVSPRNSPGEMLLNLLELGPENLYSLHLSRMVNGYAGGMQVQVVLGNAFTAGKII 120
QY 121 FAAVPPHPFVENISAAQITMCQPHVIVDVROLEPVLPLDPIRNFPHYQNTPRMRLVA 180
DB 121 FAAVPPHPFVENISAAQITMCQPHVIVDVROLEPVLPLDPIRNFPHYQNTPRMRLVA 180
QY 181 MLYTPLRAN-SEGDEVTVCVRLTRPADPDEFTFLVPPTVESKTKPFTLTLTIGELSNS 239
DB 181 MLYTPLRAN-SEGDEVTVCVRLTRPADPDEFTFLVPPTVESKTKPFTLTLTIGELSNS 239
QY 240 RFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLOQTTLQVPTQICAFRGTLSISQARAADS 299
DB 240 RFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLOQTTLQVPTQICAFRGTLSISQARAADS 299
QY 241 RFPPIEQLYTAPNETNVVQCONGRCTLDGELQGTTLQVPTQICAFRGTLSISQARAADS 299
DB 241 RFPPIEQLYTAPNETNVVQCONGRCTLDGELQGTTLQVPTQICAFRGTLSISQARAADS 299
QY 300 TDSQORARNHPLHVQVKNLQDGTQYDPTDDIPAVLGAIDFKGTVFGVASQRDVSGQOQGH 359
DB 300 TDSQORARNHPLHVQVKNLQDGTQYDPTDDIPAVLGAIDFKGTVFGVASQRDVSGQOQGH 359
QY 289 -DVTHIAGSHDYTMNLASQWNSNDPTEEPAPLGTQDFVGVKIQGLMTQT-----TREDG- 342
DB 289 -DVTHIAGSHDYTMNLASQWNSNDPTEEPAPLGTQDFVGVKIQGLMTQT-----TREDG- 342
QY 360 YATRAHEAHDITDPKYAPKLGTLILKSG-SDDENTNQPIRFTPVGM---GDNNM---RQ 412
DB 360 YATRAHEAHDITDPKYAPKLGTLILKSG-SDDENTNQPIRFTPVGM---GDNNM---RQ 412
QY 343 -STRAHKATVSTGSHVFTPKLGSVQYTTDNNDFQGTQNTKFTPVGVQDGNHNSPQQ 401
DB 343 -STRAHKATVSTGSHVFTPKLGSVQYTTDNNDFQGTQNTKFTPVGVQDGNHNSPQQ 401
QY 413 WELPDYSGRLTLNMLNLAVALSPSPFGERILFPRSTVPSAGGYIDCLIPQEWVQHFY 472
DB 413 WELPDYSGRLTLNMLNLAVALSPSPFGERILFPRSTVPSAGGYIDCLIPQEWVQHFY 472
QY 402 WVLNYSRGTGHNVLAPAVAPTFEGEQLLFFRSRVPVQVQLOGQDAFIDCLLPQEWVQHFY 461
DB 402 WVLNYSRGTGHNVLAPAVAPTFEGEQLLFFRSRVPVQVQLOGQDAFIDCLLPQEWVQHFY 461
QY 473 QEAAPSQAVLVRYNPDTRGRNIFEAKLHREGFLTVANGCNPNIVVPPNGYFRFPAWGN 532
DB 473 QEAAPSQAVLVRYNPDTRGRNIFEAKLHREGFLTVANGCNPNIVVPPNGYFRFPAWGN 532
QY 462 QEAAPQADVALLIRYNPDTRGRNIFEAKLHREGFLTVANGCNPNIVVPPNGYFRFPAWGN 521
DB 462 QEAAPQADVALLIRYNPDTRGRNIFEAKLHREGFLTVANGCNPNIVVPPNGYFRFPAWGN 521
QY 533 QFYTLAPMGSGQGRRA 549
DB 533 QFYTLAPMGSGQGRRA 549
QY 522 QFYTLAPMGSGQGRRA 538
DB 522 QFYTLAPMGSGQGRRA 538
```





```

PD 03-JUN-2003.
XX PF
XX 07-JUN-1995; 95US-00486049.
XX
XX 08-NOV-1989; 89US-00433492.
XX PR
XX 27-APR-1990; 90US-00515993.
XX PR
XX 27-AUG-1990; 90US-00573509.
XX PR
XX 06-MAY-1991; 91US-00696454.
XX
XX (BAYU ) BAYLOR COLLEGE MEDICINE.
XX PA
XX
XX Estes MK, Jiang X, Graham DY;
XX PI
XX WPI; 2003-776005/73.
XX DR
XX N-PSDB; ADC72174.
XX
XX Inducing an immune response against non-Norwalk virus agents, comprises
XX PT administering an immunogen recombinantly expressed from a cDNA from
XX PT Norwalk virus.
XX
XX Example 4; SEQ ID NO 3; 45pp; English.
XX
XX This invention relates to a novel method of inducing an immune response
XX in an individual against Norwalk virus and non-Norwalk virus agents, by
XX orally or parenterally administering an immunogen recombinantly expressed
XX or synthesised from a cDNA of Norwalk virus given in the specification.
XX CC Norwalk virus is one of the most important viral pathogens, causing acute
XX CC gastroenteritis. The invention may be used for the development of
XX CC compounds with virucidal activity or an antiviral vaccine. The present
XX CC sequence is the amino acid sequence of a protein encoded by the Norwalk
XX CC virus genome of the invention.
XX
XX Sequence 530 AA;
SQ
Query Match 39.3%; Score 1141; DB 7; Length 530;
Best Local Similarity 43.4%; Pred. No. 1.3e-96;
Matches 249; Conservative 78; Mismatches 169; Indels 78; Gaps 12;
QY 1 MKMASNDAAPSNDGAA---NLVPEAN-DEVMALEPVVGASIAAPVVVGQNIIDPWIRENF 56
DB 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSSTAVATAGQVNPIDPWIIINF 60
QY 57 VQAPQGEFTVSPRNSPGEMLLNLLELGPENLYSLRMVNGYAGGMQVQVVLGNAFTA 116
DB 61 VQAPQGEFTTSPNNTPGDVLDSLGLPHLNPFLHLSQMGVNGWNRVRLMAGNAFTA 120
QY 117 GKIIFAAVPPHPFVENISAAQITMCPHVIVDVRLQLEFVLLPLDIRNRFPHYNQENTPRM 176
DB 121 GKIIIVSCIPPGFGSHNLTIQAATLFPHVIAVDTLDPIEVLEDRVNLVFNHNDNRNQTM 180
QY 177 RLVMALYTPLRANSGE-DVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPILTIGE 235
DB 181 RLVCMLYTPLTRGGTGSFVAGVNMTCPSDFNLFPLVPPVEQKTRFPTLPLNPLSS 240
QY 236 LNSRFPAAIDMLYTDNVESTVQVQNGRCTLDGTLQGTTLQVLTQICAFRGLTISQTAR 295
DB 241 LNSRAPLTISSMGISFDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGT----- 293
QY 296 AADSTDSQPARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIIDFKG-----TVFGVASQ 348
DB 294 -----SNGTVINLTDLGTFPHPEEG-PAPIGFPLDGGCDWIMHNTQFGHSSQ 340
QY 349 RDVSGQQGHVATFAEAHTDTPDKYAPKLGITLIK-SGDDFNTNQPIRETFPVGMGD 407
DB 341 T-----QYDVTTPDTPFVPHLGGIQANGISGNY-----VG 372
QY 408 NNWRQ-----WELPDYSGRLTLNNLAPVSPSPGGRILFFRSIVPSAGCYG 455
DB 373 LSWISPPSPSGQVDLWKIPNYGSSITEATHLAPSVYPGFGVLFVFMKMPGPGAYN 432
QY 456 SGYIDCLIPQEWQHVFQEAAPSQSAVALRVYVNPDTGRNIFPAKLHREGFLTIVANGC-- 513
DB 433 ---LFCLLPQEIYIHLASEQAPTVGEAALLHVDPDGTGRNLGCEPKAYPDGLTLCVPNGAS 489

```

```

QY 514 NNPIVVPNGYFRFEAWGNQFYTLAPMGSGQRR 547
DB 490 SGPQQLPINGVFVFWVSWVSRYQLKPVGTASSAR 523

RESULT 12
AAB49701
ID AAB49701 standard; protein; 530 AA.
XX
XX AAB49701;
XX
XX 04-APR-2001 (first entry)
XX
XX Small round structured virus protein SEQ ID 2.
XX
XX Small round structured virus; SRSV; food poisoning.
XX
XX Small round structured virus.
XX
XX WO200079280-A1.
XX
XX 28-DEC-2000.
XX
XX 22-JUN-2000; 2000WO-JP004095.
XX
XX 22-JUN-1999; 99JP-00175928.
XX
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.
XX PA (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX
XX WPI; 2001-080848/09.
XX DR N-PSDB; AAF29142.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
XX PT strains for investigation of food poisoning outbreaks, contains
XX PT antibodies.
XX
XX Claim 1; Page 42-45; 84pp; Japanese.
XX
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX CC used for detecting and typing strains of SRSV in order to prevent the
XX CC spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 530 AA;
SQ
Query Match 38.9%; Score 1131; DB 4; Length 530;
Best Local Similarity 43.2%; Pred. No. 1.1e-95;
Matches 248; Conservative 77; Mismatches 171; Indels 78; Gaps 12;
QY 1 MKMASNDAAPSNDGAA---NLVPEAN-DEVMALEPVVGASIAAPVVVGQNIIDPWIRENF 56
DB 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDPAVAGSSTAVATAGQVNPIDPWIIINF 60
QY 57 VQAPQGEFTVSPRNSPGEMLLNLLELGPENLYSLRMVNGYAGGMQVQVVLGNAFTA 116
DB 61 VQAPQGEFTTSPNNTPGDVLDSLGLPHLNPFLHLSQMGVNGWNRVRLMAGNAFTA 120
QY 117 GKIIFAAVPPHPFVENISAAQITMCPHVIVDVRLQLEFVLLPLDIRNRFPHYNQENTPRM 176
DB 121 GKIIIVSCIPPGFGSHNLTIQAATLFPHVIAVDTLDPIEVLEDRVNLVFNHNDNRNQTM 180
QY 177 RLVMALYTPLRANSGE-DVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPILTIGE 235
DB 181 RLVCMLYTPLTRGGTGSFVAGVNMTCPSDFNLFPLVPPVEQKTRFPTLPLNPLSS 240
QY 236 LNSRFPAAIDMLYTDNVESTVQVQNGRCTLDGTLQGTTLQVLTQICAFRGLTISQTAR 295

```

[illegible]

RESULT 13

RESOLUTION IS  
AAB49703  
ID AAB49703 standard; protein: 544 AA.

AC AAB49703;

DT 04-APR-2001 (first entry)

Small round structured virus protein SEQ ID 4.

KW Small round structured virus; SRSV; food poisoning.

OS Small round structured virus.

PN WO200079280-A1.

PD 28-DEC-2000.

22-JUN-2000: 2000WO-JP004095.

PR 22-JUN-1999; 99JP-00175928.

PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

PA (DENK-) DENKA SEIKEN KK.

PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

DR WPI; 2001-080848/09.

DR N-PSDB; AAF29144.

PT Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.

PS Claim 1; Page 47-49; 84pp; Japanese.

This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

Sequence 544 AA;

Query Match 38.7%; Score 1124.5; DB 4; Length 544;  
Best Local Similarity 44.2%; Pred. NO. 4.5e-95;  
Matches 250; Conservative 80; Mismatches 192; Indels 43;

[illegible]

## RESULT 14

AAR57091  
 ID AAR57091 standard; protein; 530 AA.

AC AAR57091;

27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 05-OCT-1994 (first entry)

DE Small round virus SRSV/KY/89 capsid protein.

pathogen; acute gastroenteritis; food poisoning; seafood contamination;  
diagnostic assay; human calicivirus; small round virus; SRSV; KY89;  
Norwalk virus; capsid protein.

OS Small round structured virus.

PN WO9405700-A2.

PD 17-MAR-1994.

AA 07-SEP-1993: 93WO-US008447.

PR 07-SEP-1992: 92US-00941365.

XX (BAYU ) BAYLOR COLLEGE MEDICINE.

PI Matson DO. Estes MK. Jiang X. Graham DY: xx

WPI: 1994-101125/12.

Wed Jun 2 09:13:24 2004

DR N-PSDB; AAQ56832.  
 XX DNA from Norwalk and related viruses - used for preparing prods. for use  
 XX in diagnostic assays, detection and vaccines for Norwalk and related  
 XX viruses.  
 XX  
 XX Example 7; Fig 13a; 156pp; English.  
 XX  
 XX The known sequence for Norwalk virus was used to obtain the sequence of  
 XX other Norwalk-related viruses such as SRSV/KI/89, an agent from a stool  
 XX from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide  
 XX cDNA sequence includes part of the polymerase region and the capsid  
 XX region of the genome; the deduced amino acid sequences are AAR57092 and  
 XX AAR57091, respectively. Expression of fragments and derivs. of Norwalk-  
 XX related viruses permits development of diagnostic assays to detect  
 XX antibodies, antigens, viral genetic material or antivirals. (Updated on  
 XX 25-MAR-2003 to correct PN field.) (Updated on 27-AUG-2003 to correct OS  
 XX field.)  
 XX  
 XX Sequence 530 AA;  
 XX  
 XX Query Match 37.7%; Score 1097; DB 2; Length 530;  
 XX Best Local Similarity 42.8%; Pred. No. 1.6e-92; Indels 80; Gaps 14;  
 XX Matches 246; Conservative 78; Mismatches 171;  
 QY 1 MKMASNDAAPSNDGAA---NLVPEAN-DEVMALEPVVGASIAAPVVGQNIIDPWIRENF 56  
 DB 1 MMWASKDATTSSVDGASASQLVPEVNASDPLAMDVPAGSSTAVATAGQVNPIDPWIIINF 60  
 QY 57 VQAPQGEFTVSPRNSPGEMLLNLGLPELNPYLHSLRMVNGYAGGMQVQVLAGNAFTA 116  
 DB 61 VQAPQGEFTISPNTPGDVLFDLSGLPHLPFLHLSQMYNGVGNMRVIRLAGNAFTA 120  
 QY 117 GKIIFAAVPFPFVENISAAQITMCPHVIDVQLEPVLPLPDIRNRRFFHYNOENTPRM 176  
 DB 121 GKIIIVCCIPFGSGQQITIAQATLFPHVIAADVRLDPIEVPLEDVRNVLHNDRNQTM 180  
 QY 177 RLVMALYTPLRANSGE--DFTVSCRVLTRPADPFEFTFLVPPTVESKTKPFTLPLTLGE 235  
 DB 181 RLVCMLYTPLSTGGTGDSPVAGRVNMTCPSPDNFLFLVPPTVEQKTRPFTLPLNPLSS 240  
 QY 236 LNSRFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLQGTTLVPTQICAFRGTLSQSTAR 295  
 DB 241 LNSRPAFLPSGMISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGT----- 293  
 QY 296 AADSTDSQPRARNHPLHVQVKNLDGTYDDTDIPAVLGALDFKG-----TVGCVASQ 348  
 DB 294 -----SNGTVINLTDLGTPFHPFEG-PAPIGFPDLGGCDWHINMTQFGHSSQ 340  
 QY 349 RDVSGQQOQGHYATRAHEAHDITDPKYAPKLGITLIKSGSDDFNTNQPIRFTPVGMGD 407  
 DB 341 T-----QYDVTDTTDSVPHLGSIQANGISGNY-----IGV 372  
 QY 408 NNWRQ-----WELPDYSGRLTLNNMLAPAV-SPSPGQERILFRSIVPSAGGY 454  
 DB 373 LSWSPSPHSPSGSQVDLWKIPNYGSSITEATHLAPSVVSPGF-GEVLVFFMSKIPGGG- 430  
 QY 455 GSGYIDCLIPQBWQHFQEAAPSQSAVALRYVYNPDTRGNIPEAKLHREGFLTVACNG- 513  
 DB 431 --DSLPCLLIPOGYISHLASEQAPTVEGEPFLHYVDPTDRNLGEFKAYPDGFLTVCVNGA 488  
 QY 514 -NNPIVVPNGYFRFEACWGNQFTYLLAPMGSGQGR 547  
 DB 489 SSGPQOLPINGVVFVSWVSRYQLKPVGTASTAR 523  
 RESULT 15  
 ID AAB49702  
 XX AAB49702 standard; protein; 546 AA.  
 AC AAB49702;  
 XX  
 XX 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 3.  
 XX Small round structured virus; SRSV; food poisoning.  
 XX Small round structured virus.  
 XX WO200079280-A1.  
 XX 28-DEC-2000.  
 XX 22-JUN-2000; 2000WO-JP004095.  
 XX 22-JUN-1999; 99JP-00175928.  
 XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 XX (DENK-) DENKA SEIKEN KK.  
 XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 XX WPI: 2001-080848/09.  
 XX N-PSDB; AAF29143.  
 XX Kit for the detection and typing of small round-structured virus (SRSV)  
 XX strains for investigation of food poisoning outbreaks, contains  
 XX antibodies.  
 XX Claim 1; Page 45-47; 84pp; Japanese.  
 XX This invention relates to a kit for the detection and typing of small  
 XX round structured virus (SRSV) strains. The kit contains antibodies  
 XX directed against peptides represented in sequences AAB49700 - AAB49710,  
 XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 XX used for detecting and typing strains of SRSV in order to prevent the  
 XX spread of infection and to examine the epidemiology of outbreaks  
 XX  
 XX Sequence 546 AA;  
 XX  
 XX Query Match 37.5%; Score 1090.5; DB 4; Length 546;  
 XX Best Local Similarity 42.6%; Pred. No. 6.6e-92;  
 XX Matches 250; Conservative 75; Mismatches 173; Indels 89; Gaps 16;  
 QY 1 MKMASNDAAPSNDGAA---NLVPEAND-EVMALEPVVGASIAAPVVGQNIIDPWIRENF 56  
 DB 1 MMWASKDAPQSDAGSAGQIVFEVNTADPLPEPVAGPTTAVATAGQVNMIDPWVNNF 60  
 QY 57 VQAPQGEFTVSPRNSPGEMLLNLGLPELNPYLHSLRMVNGYAGGMQVQVLAGNAFTA 116  
 DB 61 VQAPQGEFTISPNTPGDILLQLGPHLPFLHLSQMYNGVGNMRVIRLAGNAFTA 120  
 QY 117 GKIIFAAVPFPFVENISAAQITMCPHVIDVQLEPVLPLPDIRNRRFFHYNOENTPRM 176  
 DB 121 GKIIIVCCVPFGTSSSLTIAQATLFPHVIAADVRLDPIEVPLEDVRNVLHTN-DNQPTM 179  
 QY 177 RLVMALYTPLRANSGE--EDVTVSCRVLTRPADPFEFTFLVPPTVESKTKPFTLPLTL 233  
 DB 180 RLVCMLYTPLRTGGSGNSDSFVWAGRVLTAPSDFSFLVPPETIEQKTRAFVNPIL 239  
 QY 234 GELNSRFPAAIDMLYTDPNESIVVQPNQGRCTLDGTLQGTTLVPTQICAFRGTLSQSTAR 293  
 DB 240 QTLNSRFPFLSQMILSPDASQVQVQFQNGRCLLDGQLLGTTPATSGQLFVRCK-INQG 298  
 QY 294 ARAADSTDSQPRARNHPLHVQVKNLDGTYDDTDIPAVLGALDF-----KG 340  
 DB 299 ARTLNLT-----VDGKPFMAFDS-PAPVGFDPFGKCDWHMRISKTPNN 341  
 QY 341 TVVFG-----VASQDVSGQOQGHYATRAHEAHDITDPKYAPKLGITLIKSGSDDFN-- 393  
 DB 342 TSSGDPMRSVSVQTNVQG-----FVPHLGSIQF---DEVFNHP 376  
 QY 394 -----TNQPIRFTPVGMGNWRQWELPDYSGRLTLNNMLAPAVSPSPGQERILF 443



Db 377 TGDYIGTIEWISQP--STPPGTDIN---LWEIPDYGSSLSQAANLAPPVPPPGFEALVY 431  
Qy 444 FRSIVPSAGGYS-GYIDCLIPQEWVQHFOEAAFPSQSAVALVRYVNPDTGRNIFEAKLH 502  
Db 432 FVSAPFGPNRRSAPNDVPCLLPQEVITHFVSEQAPTMGDAALLHYVDPTNRLGEFKLY 491  
Qy 503 REGFLTVA--NCGNPIVVPPNGYFRFEAWGNQFYTLAPMGSGQGR 547  
Db 492 PGGYLTCTVPNGVGAGPQQLPLNGVFLFVSWVSREYQLKXPGVTASTAR 538

Search completed: June 1, 2004, 13:46:01  
Job time : 49.2483 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 13.104 Seconds  
(without alignments)  
2166.837 Million cell updates/sec

Title: US-09-926-799-10  
Perfect score: 2306  
Sequence: 1 MKWASNDAPNSDGAANLVP.....GNQFYTLAPMGSGGRRRAQ 550

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/ptodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/iaa/PCTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1141	39.3	530	4	US-08-486-049-3
2	280.5	9.7	668	4	US-08-617-594A-4
3	276.5	9.5	622	4	US-09-590-020-7
4	276.5	9.5	622	4	US-09-590-020-6
5	276.5	9.5	623	4	US-09-590-020-2
6	276.5	9.5	623	4	US-09-590-020-4
7	272.5	9.4	669	4	US-08-617-594A-2
8	129.5	4.5	2318	3	US-08-091-219-24
9	129.5	4.5	2318	4	US-08-660-541-24
10	119	4.1	2647	2	US-08-583-562B-8
11	119	4.1	2647	2	US-08-779-113-8
12	109.5	3.8	1079	2	US-08-929-967-8
13	109.5	3.8	1737	4	US-09-309-572-13
14	109.5	3.8	1737	4	US-08-718-096-13
15	102.5	3.5	335	4	US-09-252-991A-28530
16	102	3.5	571	4	US-09-134-001C-3865
17	100	3.4	2232	3	US-09-091-219-25
18	100	3.4	2232	4	US-09-660-541-25
19	100	3.4	2247	3	US-09-091-219-2
20	100	3.4	2247	4	US-09-660-541-2
21	99	3.4	1091	6	5516630-2
22	99	3.4	2227	3	US-08-475-886-2
23	99	3.4	2227	3	US-08-475-886-4
24	99	3.4	2227	3	US-08-475-886-5
25	99	3.4	2227	3	US-08-397-232-2
26	99	3.4	2227	3	US-08-397-232-4
27	99	3.4	2227	3	US-09-171-387-2

28	99	3.4	2227	4	US-09-653-499-2	Sequence 2, Appli
29	99	3.4	2227	4	US-09-653-499-4	Sequence 4, Appli
30	99	3.4	2227	4	US-09-653-499-6	Sequence 6, Appli
31	99	3.4	2227	4	US-10-104-966-12	Sequence 12, Appl
32	99	3.4	2227	4	US-10-135-988-2	Sequence 2, Appli
33	99	3.4	2227	4	US-10-135-988-4	Sequence 4, Appli
34	99	3.4	2227	4	US-10-135-988-6	Sequence 6, Appli
35	98.5	3.4	1257	2	US-08-750-152A-2	Sequence 2, Appli
36	98	3.4	226	3	US-09-091-219-5	Sequence 5, Appli
37	98	3.4	226	4	US-09-660-541-5	Sequence 5, Appli
38	97.5	3.4	713	2	US-08-849-212-4	Sequence 4, Appli
39	97.5	3.4	725	1	US-08-448-170-4	Sequence 4, Appli
40	97.5	3.4	725	3	US-08-961-803-7	Sequence 7, Appli
41	97.5	3.4	832	4	US-09-489-039A-12438	Sequence 12438, A
42	97	3.3	839	1	US-08-087-016-2	Sequence 2, Appli
43	97	3.3	1269	4	US-09-645-456A-15	Sequence 15, Appl
44	97	3.3	1269	4	US-09-425-324A-15	Sequence 15, Appl
45	97	3.3	1269	4	US-09-645-791-15	Sequence 15, Appl

ALIGNMENTS

RESULT 1  
US-08-486-049-3  
; Sequence 3, Application US/08486049  
; Patent No. 6572862  
; GENERAL INFORMATION:  
; APPLICANT: Estes, Mary K  
; APPLICANT: Jiang, Xi  
; APPLICANT: Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE:  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-486-049-3

Query Match 39.3%; Score 1141; DB 4; Length 530;  
Best Local Similarity 43.4%; Pred. No. 2.46-106;  
Matches 249; Conservative 78; Mismatches 169; Indels 78; Gaps 12;  
Qy 1 MKWASNDAPNSDGA---NLVPEAN-DEVWALEPVVVGASTAAPVVGQNIIDPWIRENF 56  
Db 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDPLVAGSTAVATAGOVNPDWIIINF 60

QY 57 VOAPQGEFTVSPRNSPGEMLLNLELGPENLPYLHSLRMVNGVAGGQVQVVLGNAFTA 116  
DB 61 VOAPQGEFTVSPNPTPGDVLFDLSGLPHLAPFLHLSOMVNGVGNVRVIMLAGNAFTA 120  
QY 117 GKIIIFAAVPPHPPVENISAAQITMCPHVIVDVROLEPVLPLPDIRNRFHYNQENTPRM 176  
DB 121 GKIIIVSCIPPGGSHNLTAQATLPHVIAVDRVTLDEIPEVLEDRVNLPHNDRNQOTM 180  
QY 177 RLVMALYTPLRANSGE-DVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPFTLPILTIGE 235  
DB 181 RLVMALYTPLRANSGE-DVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPFTLPILTIGE 240  
QY 236 LNSNRPAAIDMLYTDPNESIVVQPNQGRCTLDGTLQGTQVPTQICAFRGTLSQTAR 295  
DB 241 LNSNRPAAIDMLYTDPNESIVVQPNQGRCTLDGTLQGTQVPTQICAFRGTLSQTAR 293  
QY 296 AADSTDSQPARARNHPLHVQVKNLDTQYDPTDDIPAVLGADFKGTVFGVASORDVSGQ 348  
DB 299 AADSTDSQPARARNHPLHVQVKNLDTQYDPTDDIPAVLGADFKGTVFGVASORDVSGQ 340  
QY 349 RDVSGQEQGHVATRAHEAHIDTTPDKYAPKLGTLILKS-GSDDFNTNQIRFTPVGMGD 407  
DB 341 T-----QYDVTITPDTFVPHLGSIOANGISGNY-----VGV 372  
QY 408 NNWRQ-----WELPDYSGRLTNMMLAPAVSPSPGGERILFRSIVPSAGGYG 455  
DB 373 LSWISPPSHSPSGQDVLKIPNYGSSITBATHLAPSVYPPGGEVLVFMKMPGPGAIN 432  
QY 456 SGVIDCLIPQEWVQHYQEAAPSQSAVALRVYVNDTGNRIPEAKLHREGFLTVANCG-- 513  
DB 433 ---LPCLLPQYISHLASQAPTVGEAALHVDVDTGRLNLFGEKAYPDGFLTCVPNGAS 489  
QY 514 NNPIVPPNGYFRFEAGNQFTYLPAPMGSGQGR 547  
DB 490 SGPQQLPINGVFFVMSVSRFYQLKPVGTASSAR 523  
RESULT 2  
US-09-617-594A-4  
; Sequence 4, Application US/09617594A  
; Patent No. 6541458  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.1  
; CURRENT APPLICATION NUMBER: US/09/617,594A  
; CURRENT FILING DATE: 2000-07-14  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 4  
; LENGTH: 668  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-617-594A-4  
Query Match 9.7%; Score 280.5; DB 4; Length 668;  
Best Local Similarity 25.7%; Pred. No. 3.5e-19;  
Matches 103; Conservative 56; Mismatches 147; Indels 95; Gaps 17;  
QY 11 SNDGAANLVPEANDEV--MALEPVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68  
DB 125 ADDGSVT-TPEQGLVGVVIAEPNAQMSAVADVATGKSDVSEW--EAFPSF-HTSVNWS 180  
QY 69 RNSPCEMLNLELGPENLPYLHSLRMVNGVAGGQVQVVLGNAFTAAGKIIFAAVPPHF 128  
DB 181 SETQKILFKQSLGFLNLPYLHSLAKLYVAMSGSIEVRSISGSGVFGKLAALVWPPGI 240

QY 129 -PVENISAAQITMCPHVIVDVROLEPVLPLPDIRNRFHYNQENTPRMELVAMLYTPL- 186  
DB 241 DEVQSTMLQY---PHVLFDAQVPEVIFIPDLRNSLYHL-MSDTRTSLVIMVYNDLI 296  
QY 187 -----RANSGBEDVFTVSCRVLTRPAPDEFTFLVPPTVESKTKPFTLPILTIGLSNSRF 241  
DB 297 NEYANDSNSSGCIIVTVE---TKPGDPFKHLLKPPG-----SMLTHGSIPLD 342  
QY 242 PAADIMLYTDPNES-----IVVQPNQGRCTLDGTLQGTQVPTQICAFRGTLSQTARAA 297  
DB 343 PKSSSLWLTGNRHSWIDTDFVIRP-----FVQANRHF 374  
QY 298 D---STDSPQARARNHPLHVQVKNLDTQYDPTDDIPAVLGADFKGTVFGVASORDVSGQ 354  
DB 375 DFNQETAGWSTFRFRPIITVSEKGSK-----LG-----IGVATDSIVPG- 415  
QY 355 QEGHYATRAHEAHIDTTPDKYAPKLGTLILKSGSDDFNT 394  
DB 416 -----IPDGWPDVTIPEKLTTPAGDYAITNGNNDITT 447  
RESULT 3  
US-09-590-020-7  
; Sequence 7, Application US/095900020  
; Patent No. 6355246  
; GENERAL INFORMATION:  
; APPLICANT: Kruger, John M  
; APPLICANT: Maes, Roger K  
; APPLICANT: Vilnig, Aivars  
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE  
; FILE REFERENCE: MSU 4.1-446  
; CURRENT APPLICATION NUMBER: US/09/590,020  
; CURRENT FILING DATE: 2000-06-08  
; PRIOR FILING DATE: 1999-06-10  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 7  
; LENGTH: 626  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-09-590-020-7  
Query Match 9.5%; Score 277.5; DB 4; Length 626;  
Best Local Similarity 23.7%; Pred. No. 6.2e-19;  
Matches 127; Conservative 79; Mismatches 209; Indels 121; Gaps 23;  
QY 11 SNDGAANLVPEANDEV--MALEPVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68  
DB 80 ADDGSIT-APEQGTWVGVIATFSAQMSAADMATGKSDVSEW--EAFPSF-HTSVNWS 135  
QY 69 RNSPCEMLNLELGPENLPYLHSLRMVNGVAGGQVQVVLGNAFTAAGKIIFAAVPPHF 128  
DB 136 SETQKILFKQSLGFLNLPYLHSLAKLYVAMSGSIEVRSISGSGVFGKLAALVWPPGV 195  
QY 129 -PVENISAAQITMCPHVIVDVROLEPVLPLPDIRNRFHYNQENTPRMELVAMLYTPLR 187  
DB 196 DEVQSTMLQY---PHVLFDAQVPEVIFIPDLRNSLYHL-MSDTRTSLVIMVYNDLI 251  
QY 188 ANSGEDVFTVSC--RVLTRPAPDEFTFLVPPTVESKTKPFTLPILTIGLSNSRFPAAI 245  
DB 252 NEYANDTNSSGCIIVTVEKPGDPFKHLLKPPG-----SMLTHGSPSOLIPKSS 301  
QY 246 DMLYTDPNES-----IVVQPNQGRCTLDGTLQGTQVPTQICAFRGTLSQTARAA-- 298  
DB 302 SLWIGNRHSWIDTDFIIRP-----FVQANRHFDPNQ 333  
QY 299 STDSPQARARNHPLHVQVKNLDTQYDPTDDIPAVLGADFKGTVFGVASORDVSGQSG 358  
DB 334 ETAGWSTFRFRPIISVITTEQNG-----AKLG-----IGVATDVIYVG- 370

```
Qy 359 HYATRAHEAHIDTDPKYPAPKLILKSGS--DDFNTNQPIRFTPVGMGDNWNRQWELPD 417
Db 371 -----IPDGWPDPTTGPGLIPAGDYAITNGTGNDDITTATGYDTADIIKQNTNFR----- 419
Qy 418 YSGRLTLNMLNAPAVSPSPFCERILFFRSIVPSA---GGYSGVIDC-LIPOEWV----- 468
Db 420 -----GMYICGSLORAW-GDKKISNTAFITTLTLDGDNKKINPCNTIDQSKIVVFQD 471
Qy 469 QHFYQEAAPSQSAVALVRY-----VNPDTGRNIFEAKLHREGFLTVANGNNPI 517
Db 472 AHVGKKAQTSDDTLALLGYTGIGGQAIGSDRDRVRISTLPETG----ARGGNHPI 523

RESULT 4
US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 6
; LENGTH: 622
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-6

Query Match 9.5%; Score 276.5; DB 4; Length 622;
Best Local Similarity 31.1%; Pred. No. 7.8e-19;
Matches 73; Conservative 39; Mismatches 102; Indels 21; Gaps 8;

Qy 12 NDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQQNIIDPWIRENFVQAPQGEFTVSPR 69
Db 81 DDGSIT-TPEQGTWVGVIAPSPSAQMSAAADMATGKSVDSSEW--EAFPSF-HTSVNWNSTS 136
Qy 70 NSPGEMLLNLELGPENLPYLHSLRMVNGVAGGMQVQVVLGNAFTAGKIIFAAVPPHPP 129
Db 137 ETQKILFKQSLGPLLNPLYSLHAKLYVAMSGSVFVSISGSGVFGKLAIVVPP--G 194
Qy 130 VENISAAQITMCPHVIVDVRLPEVLLPLDIRNRFHYNQENTPRMLVAMLYTPLRAN 189
Db 195 IEPVQSTSMQLQYPHVLFDAQVEPVIFAIPDLRSNLVHL-MSD'TD'TTSLVIMVYNDLINP 253
Qy 190 SGEDVFTVSC--RVLTRPADPDEFTFLVPPTVESKTKPFTLPILTILGELNSRPP 242
Db 254 YANDTNSSGCIIVTETKPGDPDFKHLKPPG-----SMLTHGSGVPSDLIP 298

RESULT 5
US-09-590-020-2
; Sequence 2, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9

Qy 12 NDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQQNIIDPWIRENFVQAPQGEFTVSPR 69
Db 81 DDGSIT-TPEQGTWVGVIAPSPSAQMSAAADMATGKSVDSSEW--EAFPSF-HTSVNWNSTS 136
Qy 70 NSPGEMLLNLELGPENLPYLHSLRMVNGVAGGMQVQVVLGNAFTAGKIIFAAVPPHPP 129
Db 137 ETQKILFKQSLGPLLNPLYSLHAKLYVAMSGSVFVSISGSGVFGKLAIVVPP--G 194
Qy 130 VENISAAQITMCPHVIVDVRLPEVLLPLDIRNRFHYNQENTPRMLVAMLYTPLRAN 189
Db 195 IEPVQSTSMQLQYPHVLFDAQVEPVIFAIPDLRSNLVHL-MSD'TD'TTSLVIMVYNDLINP 253
Qy 190 SGEDVFTVSC--RVLTRPADPDEFTFLVPPTVESKTKPFTLPILTILGELNSRPP 242
Db 254 YANDTNSSGCIIVTETKPGDPDFKHLKPPG-----SMLTHGSGVPSDLIP 298

RESULT 6
US-09-590-020-4
; Sequence 4, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match 9.5%; Score 276.5; DB 4; Length 623;
Best Local Similarity 30.6%; Pred. No. 7.8e-19;
Matches 72; Conservative 41; Mismatches 101; Indels 21; Gaps 8;

Qy 12 NDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQQNIIDPWIRENFVQAPQGEFTVSPR 69
Db 81 DDGSIT-TPEQGTWVGVIAPSPSAQMSAAADMATGKSVDSSEW--EAFPSF-HTSVNWNSTS 136
Qy 70 NSPGEMLLNLELGPENLPYLHSLRMVNGVAGGMQVQVVLGNAFTAGKIIFAAVPPHPP 129
Db 137 ETQKILFKQSLGPLLNPLYSLHAKLYVAMSGSVFVSISGSGVFGKLAIVVPP--G 194
Qy 130 VENISAAQITMCPHVIVDVRLPEVLLPLDIRNRFHYNQENTPRMLVAMLYTPLRAN 189
Db 195 IEPVQSTSMQLQYPHVLFDAQVEPVIFAIPDLRSNLVHL-MSD'TD'TTSLVIMVYNDLINP 253
Qy 190 SGEDVFTVSC--RVLTRPADPDEFTFLVPPTVESKTKPFTLPILTILGELNSRPP 242
Db 254 YANDTNSSGCIIVTETKPGDPDFKHLKPPG-----SMLTHGSGVPSDLIP 298

RESULT 7
US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
```

```
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-2

Query Match 9.5%; Score 276.5; DB 4; Length 623;
Best Local Similarity 30.6%; Pred. No. 7.8e-19;
Matches 72; Conservative 41; Mismatches 101; Indels 21; Gaps 8;

Qy 12 NDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQQNIIDPWIRENFVQAPQGEFTVSPR 69
Db 81 DDGSIT-TPEQGTWVGVIAPSPSAQMSAAADMATGKSVDSSEW--EAFPSF-HTSVNWNSTS 136
Qy 70 NSPGEMLLNLELGPENLPYLHSLRMVNGVAGGMQVQVVLGNAFTAGKIIFAAVPPHPP 129
Db 137 ETQKILFKQSLGPLLNPLYSLHAKLYVAMSGSVFVSISGSGVFGKLAIVVPP--G 194
Qy 130 VENISAAQITMCPHVIVDVRLPEVLLPLDIRNRFHYNQENTPRMLVAMLYTPLRAN 189
Db 195 IEPVQSTSMQLQYPHVLFDAQVEPVIFAIPDLRSNLVHL-MSD'TD'TTSLVIMVYNDLINP 253
Qy 190 SGEDVFTVSC--RVLTRPADPDEFTFLVPPTVESKTKPFTLPILTILGELNSRPP 242
Db 254 YANDTNSSGCIIVTETKPGDPDFKHLKPPG-----SMLTHGSGVPSDLIP 298

RESULT 6
US-09-590-020-4
; Sequence 4, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 623
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-09-590-020-4

Query Match 9.5%; Score 276.5; DB 4; Length 623;
Best Local Similarity 30.6%; Pred. No. 7.8e-19;
Matches 72; Conservative 41; Mismatches 101; Indels 21; Gaps 8;

Qy 12 NDGAANLVPEANDEV--MALEPVVVGASIAAPVVGQQNIIDPWIRENFVQAPQGEFTVSPR 69
Db 81 DDGSIT-TPEQGTWVGVIAPSPSAQMSAAADMATGKSVDSSEW--EAFPSF-HTSVNWNSTS 136
Qy 70 NSPGEMLLNLELGPENLPYLHSLRMVNGVAGGMQVQVVLGNAFTAGKIIFAAVPPHPP 129
Db 137 ETQKILFKQSLGPLLNPLYSLHAKLYVAMSGSVFVSISGSGVFGKLAIVVPP--G 194
Qy 130 VENISAAQITMCPHVIVDVRLPEVLLPLDIRNRFHYNQENTPRMLVAMLYTPLRAN 189
Db 195 IEPVQSTSMQLQYPHVLFDAQVEPVIFAIPDLRSNLVHL-MSD'TD'TTSLVIMVYNDLINP 253
Qy 190 SGEDVFTVSC--RVLTRPADPDEFTFLVPPTVESKTKPFTLPILTILGELNSRPP 242
Db 254 YANDTNSSGCIIVTETKPGDPDFKHLKPPG-----SMLTHGSGVPSDLIP 298

RESULT 7
US-09-590-020-6
; Sequence 6, Application US/09590020
; Patent No. 6355246
; GENERAL INFORMATION:
; APPLICANT: Kruger, John M
; APPLICANT: Maes, Roger K
; APPLICANT: Vilinis, Alvars
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF
; FILE REFERENCE: MSU 4.1-446
; CURRENT APPLICATION NUMBER: US/09/590,020
; CURRENT FILING DATE: 2000-06-08
; PRIOR APPLICATION NUMBER: 60/138,484
; PRIOR FILING DATE: 1999-06-10
; NUMBER OF SEQ ID NOS: 9
```

US-09-617-594A-2  
 ; Sequence 2, Application US/09617594A  
 ; Patent No. 6541458  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Audonnet, et al.  
 ; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V  
 ; FILE REFERENCE: 454313-3151.1  
 ; CURRENT APPLICATION NUMBER: US/09/617,594A  
 ; CURRENT FILING DATE: 2000-07-14  
 ; PRIOR APPLICATION NUMBER: 60/193,332  
 ; PRIOR FILING DATE: 2000-03-30  
 ; PRIOR APPLICATION NUMBER: France 00 01761  
 ; PRIOR FILING DATE: 2000-02-11  
 ; PRIOR APPLICATION NUMBER: France 99 09421  
 ; PRIOR FILING DATE: 1999-07-16  
 ; NUMBER OF SEQ ID NOS: 26  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 2  
 ; LENGTH: 669  
 ; TYPE: PRT  
 ; ORGANISM: Feline calicivirus  
 ; US-09-617-594A-2

Query Match 9.4%; Score 272.5; DB 4; Length 669;  
 Best Local Similarity 30.2%; Pred. No. 2.2e-18;  
 Matches 79; Conservative 43; Mismatches 105; Indels 35; Gaps 11;  
 QY 13 DGANL-VPEANDEV--MALEPVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSPR 69  
 DB 126 DGDSSITTEQGTGLVGGVTAESQWATAADATGKSDVSEW--BSFFSF-HTSVNMWSTS 182  
 QY 70 NSPGEMLLNLGLPELNPYLHLRMYNGYAGQMVOVVLGNATAGKIIFAAVPPHF- 128  
 DB 183 ETQKILFKQSLPLNLYLHLRMYNGYAGQMVOVVLGNATAGKIIFAAVPPHF 242  
 QY 129 PVENISAAQITMCPHIVDVRLPDLIRNRFFHYNQENTPRMLVAMLYTPL-- 186  
 DB 243 PVQSTSMLOQ---PHVLFDAQVPEVIFSPDLRSTLYLH-MSDITDTLSLVIMVNDLIN 298  
 QY 187 ----RANGSDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPFTLPILTLGELNSRPP 242  
 DB 299 PYANDSNSSCIVTVE---TKGPDKFHLKPPG-----SMLTHGSIPSDLIP 344  
 QY 243 AA-----IDMLYTPDNESIVVQP 260  
 DB 345 KSSSLWICNRYWSDITDFVIRP 366

RESULT 8  
 US-09-091-219-24  
 ; Sequence 24, Application US/09091219  
 ; Patent No. 6171592  
 ; GENERAL INFORMATION:  
 ; APPLICANT: STUDDERT, Michael J.  
 ; APPLICANT: CRABB, Brendan S.  
 ; APPLICANT: FENG, Li  
 ; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS  
 ; FILE REFERENCE: 040268/0151  
 ; CURRENT APPLICATION NUMBER: US/09/091,219  
 ; CURRENT FILING DATE: 1998-10-05  
 ; EARLIER APPLICATION NUMBER: PCT/AU96/00815  
 ; EARLIER FILING DATE: 1996-12-18  
 ; EARLIER APPLICATION NUMBER: AU PN7201  
 ; EARLIER FILING DATE: 1995-12-18  
 ; NUMBER OF SEQ ID NOS: 25  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 24  
 ; LENGTH: 2318  
 ; TYPE: PRT  
 ; ORGANISM: Foot-and-mouth disease virus  
 ; US-09-091-219-24  
 Query Match 4.5%; Score 129.5; DB 3; Length 2318;

Best Local Similarity 21.2%; Pred. No. 0.0051;  
 Matches 75; Conservative 35; Mismatches 131; Indels 113; Gaps 15;  
 QY 79 LELGPELNPYLHLRMYNGYAGQMVOVVLGNATAGKIIFAAVPPHFVVENISAAQI 138  
 DB 367 LELPTDHHKGVYSLTDSYAYMRNGWDVEVTAVGNQFNGGCLLVAMVPFLYSIQKRELYQL 426  
 QY 139 TMCCHVIVDVRLPDLIRNRFFHYNQENTPRMLVAMLYTPLRANGSDVFTV 197  
 DB 427 TLFPHQFINPRTNMTAHITVPFVGNRYDQYKVHKP--WTLVWVWVAPLTVN----- 476  
 QY 198 SCRVLTRPAPDFEFTFLVPPTVESKTKPFTLPILTLGELNSR--FPAAD-----MLYT 250  
 DB 477 -----TEGAPQIKVYANIAPT-----NVHAGEFPFSKEGIFPVACSDGYGGLVTT 521  
 QY 251 DPNES-----IVVQPN---GRCTLDGTGTTQVLPTQICAFRGTLISQTAARAADSTD 301  
 DB 522 DPKTADPVYGVKFNPPNQLPGRFT---NLLDVAEACPT-FLRFEFGVPPVYTK---TD 573  
 QY 302 SPQARNHPLHVQVKNLDG-----TOYD-----PTD----- 327  
 DB 574 SDRVLAQFDMSLAAKQMSNTFLAGLAQYTYQYSGTINLHFMFTGPTDAKARYMVAYAPPG 633  
 QY 328 -----DIPAVLGAIIDFKGTVFVGASQORDVSG 353  
 DB 634 MBPPKTPZAAAHCIHAEMWDTGLNSKFTFSIP-YLSAADYAYTASGVAETTNVQG 686

RESULT 9  
 US-09-660-541-24  
 ; Sequence 24, Application US/09660541  
 ; Patent No. 6531136  
 ; GENERAL INFORMATION:  
 ; APPLICANT: STUDDERT, Michael J.  
 ; APPLICANT: CRABB, Brendan S.  
 ; APPLICANT: FENG, Li  
 ; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS  
 ; FILE REFERENCE: 040268/0151  
 ; CURRENT APPLICATION NUMBER: US/09/660,541  
 ; CURRENT FILING DATE: 2000-09-12  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/091,219  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05  
 ; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201  
 ; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18  
 ; NUMBER OF SEQ ID NOS: 25  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 24  
 ; LENGTH: 2318  
 ; TYPE: PRT  
 ; ORGANISM: Foot-and-mouth disease virus  
 ; US-09-660-541-24

Query Match 4.5%; Score 129.5; DB 4; Length 2318;  
 Best Local Similarity 21.2%; Pred. No. 0.0051;  
 Matches 75; Conservative 35; Mismatches 131; Indels 113; Gaps 15;  
 QY 79 LELGPELNPYLHLRMYNGYAGQMVOVVLGNATAGKIIFAAVPPHFVVENISAAQI 138  
 DB 367 LELPTDHHKGVYSLTDSYAYMRNGWDVEVTAVGNQFNGGCLLVAMVPFLYSIQKRELYQL 426  
 QY 139 TMCCHVIVDVRLPDLIRNRFFHYNQENTPRMLVAMLYTPLRANGSDVFTV 197  
 DB 427 TLFPHQFINPRTNMTAHITVPFVGNRYDQYKVHKP--WTLVWVWVAPLTVN----- 476  
 QY 198 SCRVLTRPAPDFEFTFLVPPTVESKTKPFTLPILTLGELNSR--FPAAD-----MLYT 250  
 DB 477 -----TEGAPQIKVYANIAPT-----NVHAGEFPFSKEGIFPVACSDGYGGLVTT 521  
 QY 251 DPNES-----IVVQPN---GRCTLDGTGTTQVLPTQICAFRGTLISQTAARAADSTD 301  
 DB 522 DPKTADPVYGVKFNPPNQLPGRFT---NLLDVAEACPT-FLRFEFGVPPVYTK---TD 573  
 QY 302 SPQARNHPLHVQVKNLDG-----TOYD-----PTD----- 327

Db 574 SDRVLAQDMSLAAKQMSNTFLAGLAQYQYQSGTINLHFMTGTDKARYMWAYAPPG 633  
Qy 328 -----DIPAVLGAIDFKGTVFGVASQORDVSG 353  
Db 634 MEPPKTPAAAHCIHAEDWTGLNSKFTSIP-YLSAADYATASGVAETTNVQG 686

RESULT 10  
US-08-583-562B-8  
; Sequence 8, Application US/08583562B  
; Patent No. 5922570  
; GENERAL INFORMATION:  
; APPLICANT: Staunton, Donald  
; APPLICANT: Harris, Edith S.  
; TITLE OF INVENTION: Cytoplasmic Modulators of Integrin  
; TITLE OF INVENTION: Binding  
; NUMBER OF SEQUENCES: 36  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC Compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/583.562B  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Williams Jr., Joseph A.  
; REGISTRATION NUMBER: 38,659  
; REFERENCE/DOCKET NUMBER: 27866/33033  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2647 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-583-562B-8

Query Match 4.1%; Score 119; DB 2; Length 2647;  
Best Local Similarity 21.0%; Pred. No. 0.074;  
Matches 131; Conservative 64; Mismatches 219; Indels 209; Gaps 31;

Qy 48 IDPWIREN-----FVQAPQ-GEFTVSPR-----NSPGEMLLNLGLPGLNPYL--S 90  
Db 1696 VDDVVENEDGTDFITYTAPQPGKYICVRFGGSHVNPSPQVITALAGDQPSVQPLRSQ 1755

Qy 91 HLSRMVNGYACGMQVQVVLNAGNATAGKIIIPAAVPPHPPVENISAAQITMCPHIVDVQ 150  
Db 1756 QLAPQYTYAQQGQQTWA-----PERPLGVNG-----LDVTS 1787

Qy 151 LEVLLPLDPIRNRFFHYNQENTPRML-----VAMLYTPLRANGSE- 192  
Db 1788 LRPDLVIP-----FTIKKGITGEVRMPKVAQPTITDNKGTVTVRYAPSEAGLHEM 1842

Qy 193 -----DVFTVSCRVLTAPADPFTFLVPVTVSSKTKPFTLPILTGEL 236  
Db 1843 DIRYDNMHIPGSLQFYVDYVNCGRHTAYGPG-----LTHGVVNGKATFTVNTKDAGEG 1896

Qy 237 SNS---RPPAIDMLYTDPN-----SIVVQ-----PQNGRCTLDG 269  
Db 1897 GLSLAIEGFSKAEISCTDNQDGTCSVSLPVLPGDYSILVKYNBQHVPGPSFTARTVGTDD 1956

Qy 270 TLQ-----GTTQLVPTQICAFRGTLLISQFARAADSTDSP---QRARN----- 308  
Db 1957 SMRASHLKVGSAAIPINISSETDLSLTATVVPSPGREEPCLLKLRLNRHGVHSFVPRKET 2016

Qy 309 --HPLHVQKNLQGTQYDPTDDIPAVLGAIDFKGTVFGVASQORDVSGO-QFQGHVATRAH 365  
Db 2017 GEHLVHVVK-KN---GQHVASSPIPVVISQSE-----IGDASRVRVSGGGLHEGHTFEP- 2066

Qy 366 EAHITDTPKYPKGLTILIKSGSD--DFNTNQPIREFTPVGMGDNMNRQWBLPDYSGRLT 423  
Db 2067 EFLIDTRDAGY---GLSLSTIEGSKVDINTED-----LEDGTCRVTYCPTPEGNYI 2115

Qy 424 LNMNLAPAVSPSPF-----GE-----RILPFRSIVPSAGGVSGY-IDCLIPQEWQHF- 471  
Db 2116 INIKFADQHVGPSPFVKVTGEGRVKESITRRRRAPSVANVGSCHDSLKLIPEISIQDWT 2175

Qy 472 YQEAAPS-----QSAVALRVYVNPDTGNIFEAKLHREGFLTVANGCNPPIVVP 520  
Db 2176 AQVTPSGKTHEAIEVEGENHTYCIRFPVPAEMGHTHTVSVKYKGQHV----- 2221

Qy 521 PNGYFRFEAMGNQFYTLAPMGSG 543  
Db 2222 PGSPFQF-----TVGPLGEG 2236

RESULT 11  
US-08-779-113-8  
; Sequence 8, Application US/08779113  
; Patent No. 5948991  
; GENERAL INFORMATION:  
; APPLICANT: Staunton, Donald E.  
; APPLICANT: Harris, Edith S.  
; TITLE OF INVENTION: Cytoplasmic Modulators of Integrin  
; TITLE OF INVENTION: Binding  
; NUMBER OF SEQUENCES: 65  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Marshall, O'Toole, Gerstein, Murray & Borun  
; STREET: 233 South Wacker Drive, 6300 Sears Tower  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: United States of America  
; ZIP: 60606  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/779.113  
; FILING DATE:  
; CLASSIFICATION: 536  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Greta E. No. 5948991and  
; REGISTRATION NUMBER: 35,302  
; REFERENCE/DOCKET NUMBER: 27866/33773  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 312-474-6300  
; TELEFAX: 312-474-0448  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 2647 amino acids  
; TYPE: amino acid  
; TOPOLOGY: not relevant  
; MOLECULE TYPE: protein  
US-08-779-113-8

Query Match 4.1%; Score 119; DB 2; Length 2647;  
Best Local Similarity 21.0%; Pred. No. 0.074;  
Matches 131; Conservative 64; Mismatches 219; Indels 209; Gaps 31;

Qy 48 IDPWIREN-----FVQAPQ-GEFTVSPR-----NSPGEMLLNLGLPGLNPYL--S 90  
Db 1696 VDDVVENEDGTDFITYTAPQPGKYICVRFGGSHVNPSPQVITALAGDQPSVQPLRSQ 1755

QY	91	HLRMYNGVAGGMQVQVVLVAGNAFTAGKII	FAAVPHPEPVENISAAQITMCPHVI	DYVRQ	150
DB	1756	QLAPQVTVAAQQQOTWA	-----PERPLVGVG-	-----LDVTS	1787
QY	151	LEPVLPLPLDIRRPHYNOENTPRML	-----VAMLYTPLRANSGE-	192	
DB	1788	LRPFDLVIP	-----FTIKGEITGEVRMP	SGKVAQPTITDNKDGTVTVRYAPSEAGLHEM	1842
QY	193	-----DVFVSCRVLTRAPADPEFTFLVP	PVESKTKPFTLPILTLGEL	236	
DB	1843	DIRYDNMHPGSPLOQFYVDVYVNCGHVTA	YCGP-----LTHGVVVKPATFTVTNKDAGEG	1896	
QY	237	SNS	---RFPAAIDMLYTDPE	-----SIVVQ	269
DB	1897	GUSLAIEGFSKABISCTDNQDGTCSVYLP	VLPGDYSILVYKNEQHVPGSPFTARVTGDD	1956	
QY	270	TLQ	-----GTTQLVPTQICAFRGTLISQ	TAAADTDSF---ORARN	308
DB	1957	SMRMSHLKVGSAADIPINISETDLSLTAT	VPPSGREEPCLLKRLNRHGVGISFVKET	2016	
QY	309	--HPLHVQVKNLQGTQVDPD	DDIPAVLGAIDFKGTVFGVASQ	RDSVQ--QEQGHYATRAH	365
DB	2017	GEHLVHVK-KN	---GQHVASSPIPVISQSE	-----IGDASRVVSQGLHEGHTFEPA-	2066
QY	366	EAHIDTTPDKYAPKGLTILIKGSD	--DFNTNQPIRFTPVGMGDNWRQWELPDYSGRLT	423	
DB	2067	EFIIDTRDAGY	--GLSLSIEGSPSKVDINTED	-----LEDGTCRVTYCPTPEPNI	2115
QY	424	LNNMLAPAVSPSP	-----GE---RILFRSIVPSAGGYG	SY-IDCLIPQEWVQHF-	471
DB	2116	INTKAFDQHVGPSPFVKVTEGRVKES	ITRRRAPSVANVGHCDLSLKIP	PEISIQDMT	2175
QY	472	YQEAAPS	-----QSAVALRVYVNDPTGTNR	NIPEAKLHREGFITVANCNPIVVP	520
DB	2176	AQVTPSPGKTHEABIVEGNHTY	CIRFVPAEMGHTTVSVYKGGHV	-----	2221
QY	521	PNGYFRFEANGNQFYTLAPMSG	543		
DB	2222	PGSPFQF	-----TVGPLGEG	2236	

RESULT 12

US-08-929-967-8

; Sequence 8, Application US/08929967

; Patent No. 5891637

; GENERAL INFORMATION:

; APPLICANT: Rupert, Siegfried J.W.

; TITLE OF INVENTION: Construction of Full-length cDNA Libraries

; NUMBER OF SEQUENCES: 16

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Genentech, Inc.

; STREET: 1 DNA Way

; CITY: South San Francisco

; STATE: California

; COUNTRY: USA

; ZIP: 94080

; COMPUTER READABLE FORM:

; MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk

; COMPUTER: IBM PC Compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WinPatIn (Genentech)

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/929,967

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Dreger, Ginger R.

; REGISTRATION NUMBER: 33,055

; REFERENCE/DOCKET NUMBER: P1035R1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 650/225-3216

; TELEFAX: 650/952-9881



Query Match	3.8%	Score 109.5	DB 4	Length 1737
Best local Similarity	19.3%	Pred. No. 0.33	229	Indels 181
Matches 116	Conservative	74	Mismatches	Gaps 25
Qy	24	DEVNALPFWVGASTAAPVVGQNIIDPWIRENFVQAQPGSEFTVSPRNSGEMLINTLGLP	83	
Db	997	DQQKAYQEIQAALLTAPALGLPDLTKFF--ELFVDEKQGY-----AKGVLTQKLG	1045	
Qy	84	ELNPYLHLSRWNGYAGG---MQQVVLAGNAFTAGKIIFAAPPHPPVENISAAQIT	139	
Db	1046	WRRE-VAYLSKLLDPVAAAGMPCCPVMAAATVLTQDAGL-----TWGQPLVI	1092	
Qy	140	MCPHVIVDVROLEPVLPLDP--LRN-RFHYNQENTPRMLVAMLYTPLRANSGEQVFT	196	
Db	1093	LAPHAV-----EALVKQPPDRWLSNARMTHYQ-----ALLDTRVQGPVVAL	1136	
Qy	197	VSCRVLTRPADDFETFLVPPTVESKTKPFTLPILTGLSELSNRPFAADMLYTDPNESI	256	
Db	1137	NPATLLPLPEGLQHNCNLDILAAEHGRP-----DLTDQPLPADHTWTYDGSLL	1187	
Qy	257	VVQPQNGRCTLDGLTQGTOLV-----PTQICAFRGTLI--SQARAADS-----TDSP	303	
Db	1188	----QEQRKAGAAVTTETETEVIAKALPAGTSAQRAELIALTQALQWAEKGLNVYDSR	1243	
Qy	304	QRARNHPLHVQV-----KNLDGTQYDPTDDIPAVLGA	336	
Db	1244	YAFATAIHGEIYIRRRGLLTSEGKEIKNKDBEILALLKALPLKRLSIHCPGHQKHSAE	1303	
Qy	337	-----DFKGTVFVGASQRDVSGQQQGHVATRAEAHIDTDPKYAPKGLTLLIKSGS	389	
Db	1304	ARGNEMADQAARKAAITETPTDTSLLIENSPPYTSEHPHYVTVDIKOLTKGLIYDTKK	1363	
Qy	390	DDFNWNOPI---RPTFGMGDNWRQWELPDYSGLT-----LNNMLAPAVSPF----	436	
Db	1364	YWVYQGRVMPDQFT-----FELDLFHLQTLHLSFKMKALLERSHSPYYMLNR	1412	
Qy	437	-----PCERI-----LFPRSIVPSAGGVGSGYIDC	461	
Db	1413	DRTLKNITETKACAQVNASKSAVKQTRVRGHRPGTHWEIDFTEIKP--GLVYKYLIV	1470	
Qy	462	LIP--QEWVQHIFYQEAAPSOASAVALRVYNPDTGRNITFEAKLHREGFLTVCANGNP	519	
Db	1471	FIDTFSGWI-----EAPPTKETAKV-----VTKLLEBIFFPFMGPOVLGTDNGAPV	1519	

Best Local Similarity 19.3%; Pred: No. 0.33;	
Matches 116; Conservative 74; Mismatches 229; Indels 181; Gaps 25;	
Qy	24 DEVMALPVPVVGASIAAPVVGQQNIIDPWIRENFVQAPQGEFTVSPRNSPGEMLLNLELGP 83
Db	197 DQKAVQEIQAALLTAPALGPLDLTKPF--ELFVDEKQGY-----AKGVLTKQLGP 1045
Qy	84 ELPNYSLSRMVYAGG---MQVVVLVAGNAFTAGKIIIPAAPPVPPFVENISAAQIT 139
Db	1046 WRRP-VAYLSKKLDPVAAWPPCLLR:::     -----TWGQPLVI 1092
Qy	140 MCRHVIYDVRQLBPVLLPLPD--IRN-RFPHYNQENTPRMLVAMLYTPLRANSGEDVFT 196
Db	1093 LAHAV-----EALVKQPPDRMLSNAEMTHYQ-----ALLDTRVQFGPVVAL 1136
Qy	197 VSCRVLTRPAPDPEFTFLVPPTVESKTKPTFLILTLGELNSRFPAAIDMLTYDNPESI 256
Db	1137 NPATLLPLPBEGLQHNCCLDLABAHGTRP-----DLTDQPLPDADHTWYTDGSSLL 1187
Qy	257 VVQPQNGRCRLDGTLOQTTQLV-----PTQICAFRGTLLI--SQTARAADS-----TDSP 303
Db	1188 ----QEQORKAGAAVTETEVIWAKALPAGTSQAQRAELIALTQALWAECKLNVYTDSR 1243
Qy	304 QRARNHPLHVQV-----KNLDTQVDPDIDIPAVLGAI-----336
Db	1244 YAFATAIHGEIYRRRGLLTSEGEIKNKDEIALLALKALPLKRLSIIHCPHQKGHSAE 1303
Qy	337 -----DPKGTVFGVASQRDVSQGQEOGHYATRAEAHIDTTPKYAPKGLTILIKSGS 389
Db	1304 ARGNRMAQARAAATEPDTSTLLIENSSPYTSEHFHYVTVDIKDLTKLGAIDYKTKK 1363
Qy	390 DDFNTNQPI---RPTPVGMGDNNRWQELPDYSGRLT-----LNNMLAPAVGSP--436
Db	1364 YWYQYQKPVMPDQFT-----PELLDFLQLHTLFSQKALLERSHSPYYMLNR 1412
Qy	437 -----PGERI-----LPFRSIVPSAGGVSQGYIDC 461
Db	1413 DRLLKNITETKCAQVNAASKSAVQKTRVGRHPRGTHWEIDTEIKP--GLGYKXLLV 1470
Qy	462 LIP--QEWQHFYQEAAPGSAVALRYVNPDTGRNIFEAKLHREGFLTIVANGNNPIV 519
Db	1471 FIDTFSGWI-----EAPPTKKEATKV-----VTKKLBEIPFRGQVPLGTGNGPAFV 1519

QY	282	ICA	F	G	T	L	I	S	O	T	A	R	A	A	D	S	T	P	O	R	A	R	N	H	P	L	V	H	V	Q	N	L	D	G	T	Q	Y	-----	D	T	D	I	P	330													
DB	81	V	L	A	L	Q	A	L	-----	R	E	H	I	L	V	F	R	Q	L	D	E	Q	L	R	F	A	T	L	E	G	S	V	F	Q	P	A	D	I	124																		
QY	331	A	V	L	G	A	I	D	F	K	-----	G	T	V	G	V	A	S	O	R	D	V	S	Q	O	Q	G	H	Y	A	T	A	R	A	E	A	H	I	D	T	D	P	K	V	A	K	L	G	I	L	I	K	S	G	S	389	
DB	125	V	L	S	G	G	D	K	V	D	I	-----	T	G	D	G	E	L	G	N	F	A	L	P	A	H	I	D	H	Q	W	T	P	V	P	S	G	S	G	L	V	A	L	E	V	P	S	G	180								
QY	390	D	E	F	N	T	Q	I	P	R	I	-----	F	T	V	G	M	G	N	N	R	Q	W	E	L	P	D	Y	G	S	R	L	T	L	N	N	L	A	P	A	V	S	P	P	G	E	R	I	L	F	F	R	S	I	V	P	449
DB	181	G	E	-----	T	R	F	T	-----	N	L	A	R	A	V	E	S	L	D	E	A	T	R	R	E	I	D	-----	G	L	R	I	N	P	F	I	R	-----	216																		
QY	450	-	S	A	G	G	Y	G	S	Y	I	D	C	L	I	P	O	E	W	H	F	Q	E	A	P	S	O	-----	S	A	V	A	L	V	R	V	Y	N	P	D	T	G	R	N	I	E	A	K	L	H	R	E	504				
DB	217	L	R	E	G	Y	G	G	F	A	T	Y	R	T	-----	D	I	E	P	O	G	S	H	P	L	R	-----	T	H	P	E	S	G	R	V	I	F	L	S	A	H	T	E	263													

Search completed: June 1, 2004, 13:58:09  
Job time : 16.104 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 35.1594 Seconds  
(without alignments)  
4368.312 Million cell updates/sec

Title: US-09-926-799-10  
Perfect score: 2906  
Sequence: 1 MKMASNDAPNSDGAANLVP.....GNQPYTLAPMGSGGRRRAQ 550

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/1/pubpaa/FCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/1/pubpaa/FCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1141	39.3	530	14	US-10-314-739-3
2	280.5	9.7	547	12	US-10-670-695-4
3	280.5	9.7	668	14	US-10-209-507-4
4	280.5	9.7	671	12	US-10-670-695-2
5	272.5	9.4	669	14	US-10-209-507-2
6	124.5	4.3	1147	15	US-10-327-481A-38
7	121	4.2	1344	9	US-09-738-626-6888
8	119	4.1	2609	14	US-10-043-487-583
9	119	4.1	2647	12	US-10-231-956A-87
10	111.5	3.8	1199	16	US-10-677-558-2
11	104.5	3.6	944	14	US-10-174-677-101
12	104	3.6	368	14	US-10-156-761-7936
13	103.5	3.6	394	15	US-10-104-047-3560
14	103	3.5	3317	15	US-10-080-334-200
15	102	3.5	1167	12	US-10-282-122A-62618

16	102	3.5	1184	12	US-10-282-122A-64572	Sequence 64572, A
17	101	3.5	563	12	US-10-425-114-59009	Sequence 59009, A
18	101	3.5	572	12	US-10-425-114-46359	Sequence 46359, A
19	101	3.5	723	12	US-10-425-114-38441	Sequence 38441, A
20	101	3.5	815	12	US-10-425-114-39399	Sequence 39399, A
21	101	3.5	815	12	US-10-425-114-62649	Sequence 62649, A
22	100.5	3.5	410	12	US-10-282-122A-48886	Sequence 48886, A
23	100.5	3.5	1307	15	US-10-104-047-2438	Sequence 2438, Ap
24	100.5	3.5	6620	15	US-10-080-334-290	Sequence 290, App
25	100	3.4	1794	10	US-09-965-738-299	Sequence 299, App
26	100	3.4	1799	10	US-09-965-738-149	Sequence 149, App
27	100	3.4	1821	10	US-09-965-738-82	Sequence 82, Appl
28	100	3.4	2234	12	US-10-612-090-20	Sequence 20, Appl
29	100	3.4	5877	14	US-10-142-515-11	Sequence 11, Appl
30	100	3.4	5935	14	US-10-243-243A-8	Sequence 8, Appli
31	100	3.4	11721	10	US-09-965-738-162	Sequence 162, App
32	99.5	3.4	534	12	US-10-282-122A-64860	Sequence 64860, A
33	99.5	3.4	1257	9	US-09-738-626-4750	Sequence 4750, Ap
34	99.5	3.4	2435	12	US-10-282-122A-47453	Sequence 47453, A
35	99	3.4	836	14	US-10-272-459-40	Sequence 40, Appl
36	99	3.4	980	14	US-10-272-459-41	Sequence 41, Appl
37	99	3.4	1382	14	US-10-010-160-18	Sequence 18, Appl
38	99	3.4	2227	9	US-09-929-955-12	Sequence 12, Appl
39	99	3.4	2227	13	US-10-104-966-12	Sequence 12, Appl
40	99	3.4	2227	13	US-10-135-988-2	Sequence 2, Appli
41	99	3.4	2227	13	US-10-135-988-4	Sequence 4, Appli
42	99	3.4	2227	13	US-10-135-988-6	Sequence 6, Appli
43	99	3.4	2227	16	US-10-719-619-12	Sequence 12, Appl
44	97.5	3.4	713	15	US-10-369-493-708	Sequence 708, App
45	97.5	3.4	2365	14	US-10-156-761-14890	Sequence 14890, A

ALIGNMENTS

RESULT 1

US-10-314-739-3  
; Sequence 3, Application US/10314739  
; Publication No. US20030129588A1  
; GENERAL INFORMATION:  
; APPLICANT: Bates, Mary K  
; Jiang, Xi  
; Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; Characterize No. US20030129588A1walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE: <Unknown>  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION NUMBER: US/10/314,739  
; APPLICATION NUMBER: US/10/314,739  
; FILING DATE: 09-Dec-2002  
; CLASSIFICATION: <Unknown>  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 530 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-10-314-739-3

Query Match 39.3%; Score 1141; DB 14; Length 530;  
Best Local Similarity 43.4%; Pred. No. 2.1e-103;  
Matches 249; Conservative 78; Mismatches 169; Indels 78; Gaps 12;

QY 1 MKMASNDAPSNDGAA---NLVPEAN-DEVMALEPVVGASIAAPVVGQNIIDPWIRENF 56  
DB 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLANDPVAGSSATAVAGQVNPIDPWIINF 60

QY 57 VQAPQGEFTVSPRNSPGEMLLNLGLPELNPYLSHRMYNGYAGMOVVVLGNAFTA 116  
DB 61 VQAPQGEFTISPNTDGVLDLSLGPLNPLFLLHLSQWYNGWGNMRVIRMLGNAFTA 120

QY 117 GKIIFAAVPHPPVENISAAQITMCCHVIVDVRQLEPVLPLPDIRNRFFHYNQENTPRM 176  
DB 121 GKIIIVSCIPPGSHNLTAQATLPHVIVADVTLDPIEVPLEDVNRVLFHNNDRNQOTM 180

QY 177 RLVMALYTPLRANSGB-DVFTVSCRVLTRPAPDFEFTFLVPPTVSKTKPFTLPILTGL 235  
DB 181 RLVMALYTPLRANSGB-DVFTVSCRVLTRPAPDFEFTFLVPPTVSKTKPFTLPILTGL 240

QY 236 LNSRPPAAIDMLYTPNESIVVQPNRGCTLDGTLQGTQVLTQICAFRGTLISQATAR 295  
DB 241 LNSRAPLPISSMGISPDNVQSVQFQNGRCTLDGRLVGTTPVLSHVAKIRGT 293

QY 296 AADSTDSQPARNHLPHVQVKNLDGTQDPTDDIPAVLGAIDFKG 348  
DB 294 -----SNGTVINLTDLGTFPFHFEF-PAPIGPDLGGCDWHNNMTQFHSSQ 340

QY 349 RDVSGQOEGHYATRAHEAHDITDPKYAPKLTILIKS-GSDFTNQPRTPTVGMCD 407  
DB 341 T-----QYDVTDTPTDFVPHLGSIQANGISGNY 372

QY 408 NNWFO-----WELPDYSGRLLTNMNLAPAVSPFGERILFFRSIVPSAGGYG 455  
DB 373 LSWISPPSPHSGSVQDLWKIPNYGSSITEATHLAPSVYPPGFEVLVFFMSKMPGAYN 432

QY 456 SGYIDCLIPQWVQHFOEAPQSQSAVALVRYNPDTRGRNIFPEAKLHREGFLTVANGC-- 513  
DB 433 ---LPCLLPQEVISHLASEQAPTVEAALLHYVDPTGRNLGFEKAYPDGFLTCVPGNGAS 489

QY 514 NNPIVPPNGYFRFANGNQFTLAPWGSGQGR 547  
DB 490 SGPOQLPINGVFVFSWSRFRYQLKPVGTASSAR 523

RESULT 2  
US-10-670-695-4  
Sequence 4, Application US/10670695  
Publication No. US20040058316A1  
GENERAL INFORMATION:  
APPLICANT: Jensen, Wayne A.  
APPLICANT: Lappin, Michael R.  
APPLICANT: Rosen, David K.  
APPLICANT: Andrews, Janet S.  
TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE  
FILE REFERENCE: DI-9-1  
CURRENT APPLICATION NUMBER: US/10/670,695  
CURRENT FILING DATE: 2003-09-25  
PRIOR APPLICATION NUMBER: 09/521,738  
PRIOR FILING DATE: 2000-03-09  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 4

LENGTH: 547  
TYPE: PRT  
ORGANISM: Feline calicivirus  
US-10-670-695-4

Query Match 9.7%; Score 280.5; DB 12; Length 547;  
Best Local Similarity 24.1%; Pred. No. 2.6e-18;  
Matches 130; Conservative 74; Mismatches 207; Indels 129; Gaps 24;

QY 11 SNDGAANLVPEANDEV--MALEPVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68  
DB 1 ADDGSIT-APEQGTWVGVIAPESQAMSTADATGKSVDSW--EAFSP-HTSVMWST 56

QY 69 RNSPGMLNLNLGLPELNPYLSHRMYNGYAGMOVVVLGNAFTAAGKIIFAAPVPHF 128  
DB 57 SETQKILFKQSGLPLNPYLEHLAKLYVWAGSIEVRSISGSGVFGKLAIVVPGV 116

QY 129 -PVENISAAQITMCCHVIVDVRQLEPVLPLPDIRNRFFHYNQENTPRMLVAMLYTLP- 186  
DB 117 DPVQSTMLQY---PHVLFDAQVEVIFCLPDLRSTLYHL-MSDITDTLSLVMYNDLI 172

QY 187 -----RANGSEDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPFTLPILTGLGELS 241  
DB 173 NPVANDANSGCCIVTVE-----TKPGDFEFHLLKPPG-----SMLTHGSIPLS 218

QY 242 PAA-----IDMLYTPNESIVVQPNRGCTLDGTLQGTQVLTQICAFRGTLISQATAR 297  
DB 219 PKTSSLWIGNRYWSDITDFVIRP-----FVQANRHF 250

QY 298 D-----STDSQPARNHLPHVQVKNLDGTQDPTDDIPAVLGAIDFKGTVFGVASQRDVSG 354  
DB 251 DFNQETAGWSTPRFRPISVTITEQNG-----AKLG-----IGVATDYIVPG- 291

QY 355 QEQGHYATRAHEAHDITDPKYAPKLTILIKS-GSDFTNQPRTPTVGMCDNNRWQ 413  
DB 292 -----IPDGWPDITIPGELIPAGDYAITNGTGNDITATGATDADIKNNTNFR-- 340

QY 414 ELPDYSGRLLTNMNLAPAVSPFGERILFFRSIVPSA---GGVSGYIDC-LIPOEWV- 468  
DB 341 -----GMVYCSQORAW-GDKKISNTAFTTATLDGNNKINFCNTIDQSKIV 388

QY 469 ---QHFOEAPQSQSAVALVRY-----VNPDTGRNIFPEAKLHREGFLTVANGC 517  
DB 389 VFQDNHVKKKAQTSDDTLALLGYTGIGEAIGSDRDRWRVIRSTLPETG----ARGGNHPI 444

RESULT 3  
US-10-209-507-4  
Sequence 4, Application US/10209507  
Publication No. US20030109033A1  
GENERAL INFORMATION:  
APPLICANT: Audonnet, et al.  
TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V  
FILE REFERENCE: 454313-3151.2  
CURRENT APPLICATION NUMBER: US/10/209,507  
CURRENT FILING DATE: 2002-11-12  
PRIOR APPLICATION NUMBER: 09/617,594  
PRIOR FILING DATE: 2000-07-14  
PRIOR APPLICATION NUMBER: 60/193,332  
PRIOR FILING DATE: 2000-03-30  
PRIOR APPLICATION NUMBER: France 00 01761  
PRIOR FILING DATE: 2000-02-11  
PRIOR APPLICATION NUMBER: France 99 09421  
PRIOR FILING DATE: 1999-07-16  
NUMBER OF SEQ ID NOS: 26  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 4  
LENGTH: 668  
TYPE: PRT  
ORGANISM: Feline calicivirus  
US-10-209-507-4

Query Match 9.7%; Score 280.5; DB 14; Length 668;

Best Local Similarity    25.7%, Pred. No. 3.6e-18;  
Matches 103; Conservative    56; Mismatches 147; Indels         95; Gaps      17;

Qy                  11   SNDGAANLVPEANDV--MALEPVVGASIAAPVVVGOQNIIIDPWIRENFVQAPOQEFTVSP    68  
Db                  :::| | | | | | | | | | | | | | | | | | | | | | | | | |  
125   ADDGSVT-TPEOGTLVGVIAPNAQSADVAATGKSVDEW--EAFPSF-HTSVNWSI    180

Qy                  69   RNSPGEMLLNLEGPELNPYLSHLSRWYNGVAGMQOVVLGNAGFAITAGKIIFAAVPPHF    128  
Db                  :::| | | | | | | | | | | | | | | | | | | | | | | | | |  
181   SETQGKLTFKSGLGPLNLTHLAKEYAWSSIEVRFSISGVFGCKLAATIVVPPI    240

Qy                  129 -PVENISAQQITMCPHVIDVROLEPYLLPDINRFFHYNQENTPRMBRLVMALYTPL-    186  
Db                  ||:||||| | | | | | | | | | | | | | | | | | | | | | | | | | |  
241   DPVOSTSMQLY---PHVLFARQVEPFIFTPDLRNSLYHL-MSDTDTTSILVIMYNLDLI    296

Qy                  .        ---RANGSEDVFVSRCVLTRPADPDFETFLVPPTVESKTGFPTLPILTGLSLSNR    241  
Db                  :||| | | | | | | | | | | | | | | | | | | | | | | | | | |  
297   NPAYDNNSSGCIVTE----TKGPDPFKFHLLKPPG-----SMLTHSGIPS DLI    342

Qy                  242   PAADMPLYTDPNES-----IVQPONGRCTLDGTLOGTTQLTVPTOICAFRGTLISQTAAA    297  
Db                  :||| | | | | | | | | | | | | | | | | | | | | | | | | | |  
343   PKSSLWGIRHSMDINDFIKP-----FVFPQANRH    374

Qy                  298   D---STDSPQRARNHPLHVOKNLDGTQYDPTDDIPAVLGAI DFKGTVFGVASORDYSYGQ    354  
Db                  :||| | | | | | | | | | | | | | | | | | | | | | | | | | |  
375   DFNOETAGNTGTPRPRTTIIVSEKGGSK-----IGVATDSIVPG-    415

Qy                  355   QEOGHYATRAREAHIDTDF-KYAPKLTGLIKSGSDDFNT    394  
Db                  :||| | | | | | | | | | | | | | | | | | | | | | | | | | |  
416   -----IPDGWPDTITEPKLTPAGDYAITNGNDNIIT    447

RESULT 4  
US-10-670-695-2  
Sequence 2, Application US/10670695  
Publication No. US20040058316A1  
GENERAL INFORMATION:  
Applicant: Jensen, Wayne A.  
Applicant: Leppin, Michael R.  
Applicant: Rosen, David K.  
Applicant: Andrews, Janet S.  
Title Of Invention: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE  
FILE REFERENCE: DI-9-1  
CURRENT APPLICATION NUMBER: US/10/670,695  
PRIOR FILING DATE: 2003-09-25  
PRIOR APPLICATION NUMBER: 09/521,738  
NUMBER OF SEQ ID NOS: 36  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 2  
LENGTH: 671  
TYPE: PRT  
ORGANISM: Feline calicivirus  
US-10-670-695-2

```

297 NPYANDANSGCCIVTVE-----TKRGPDPKFHLKPPG-----SMLTHGS:PSDLI 342
242 PAA-----IDMLYTDPNESIVVOPQNGRCTLDTGLTQGTQLVPTQICAPRGTLISOTARAA 297
343 PKTSSLWIGNRYSWDITDFVIRP-----FVFOANRHP 374
298 D---STDSPQARNNHPLHVQKNLDGTYDPTDDIPAVLGAIDFKGTGVGVASORDVSGQ 354
375 DFNQETAGWSTPRPRPISVTITEQNG-----AKLG-----IGVATDVIYPG- 415
355 QEOGHYATRAHAHIDTTPKYAPKLGTLIKSGS-DDFNTNQPIRFTPVGMGDDNNRWQW 413
416 -----IPDQGWPDITIPGELIPAGDYAITNGTGDNDITATGYDGTADIKNNTNFR-- 464
414 ELPDYSGRLLTNMNLAPVSPSPGGERILPFRSIVPSA---GGVSGGYIDC-LIPQEMV- 468
465 -----GMVYCSLQRAW-GDKKISNTAFITATLDGDDNNKINPCNTIDQSKIV 512
469 ---QHFEYQBAAPSQSAVALVRY-----VNPDTGRNIFBAKLHREGFLTVANGCNPPI 517
513 VFQDNHVKKRAQTSDDTLALLGYTGIGEAIGSDRDRVVRISTLPETG---ARGGNHPI 568

RESULT 5
US-10-209-507-2
; Sequence 2, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECO
; FILE REFERENCE: 454313-3151.2
; CURRENT APPLICATION NUMBER: US/10/209,507
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR APPLICATION NUMBER: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 669
; TYPE: PRT
; ORGANISM: Feline calicivirus
US-10-209-507-2

```

Query Match	9.7%	Score 280.5	DB 12	Length 671
Best Local Similarity	24.1%	Pred.No. 3.6e-18		
Matches 130; Conservative	74	Mismatches 207	Indels 129	Gaps 24

  

Qy	11	SDNGAANLVPEANDEV--MALEPVPVGGASIAAPVVGQNIIDPMIRENFVQAPQGEFTVSP	68
Db	125	ADDSGIT-APBQGTWGCVTAAEASQAMSTAAADWATGKSDSEW--EAFFSF-HTSVNWMT	180
Qy	69	RNSPGEMLLNLELGPENLYSHLSRMYNGYAGGMQVGVVLGAGNAFTAGKIIIFAAVPPHF	128
Db	181	SETQCKILFKQSLGPLLNPLYLEHLAKLYAWSSGSEIEVFSITSGSGVFGGKLAATVVPVG	240
Qy	129	-PVENISAAQITMCPHVIVDVQLEPVLPLPDIFRNFHFYVNOENTPRMLVAMLYTPL-	186
Db	241	DPVQSTSMQLQY---PHVLFQARQEPVIFCUPDLURSTLYHL-KSDTDTTSLVTVMYNDLI	296
Qy	187	-----RANSGEDVFTVSCRVLTPAPDFEFTFLVPPTVESKTFPTLLTILGELSRSR	241

Wed Jun 2 09:13:24 2004

RESULT 6  
US-10-327-481A-38  
; Sequence 38, Application US/10327481A  
; Publication No. US20040001964A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Andrew M.Q.  
; APPLICANT: Audonnet, Jean-Christophe F.  
; APPLICANT: Lombard, Michel F.A.  
; TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease  
; FILE REFERENCE: 454313-3178  
; CURRENT APPLICATION NUMBER: US/10/327,481A  
; CURRENT FILING DATE: 2002-12-20  
; PRIOR APPLICATION NUMBER: PCT/FR01/02042  
; PRIOR FILING DATE: 2001-06-27  
; PRIOR APPLICATION NUMBER: FR 0008437  
; PRIOR FILING DATE: 2000-06-29  
; NUMBER OF SEQ ID NOS: 41  
; SOFTWARE: Patent in ver. 2.1  
; SEQ ID NO 38  
; LENGTH: 1147  
; TYPE: PRT  
; ORGANISM: Foot-and-mouth disease virus  
US-10-327-481A-38

Query Match 4.3%; Score 124.5; DB 15; Length 1147;  
Best Local Similarity 19.3%; Pred. No. 0.023;  
Matches 67; Conservative 41; Mismatches 158; Indels 81; Gaps 6;

Qy 70 NSPGEMLNLELGPENLPYLSRMVYAGQVQVVLGNAFTAGKIIFAVPPHP 129  
Db 158 DKFPGYLTLELPTDHHGVGHLVDSYAYWRNGWDEVSAGVQFNGGCLLVAMPEWKA 217  
Qy 130 VENISAAQITWCPHVIVDVQLEPVLPLDIRNRPYHYNQENTPRMLVAMLYTLPLRAN 189  
Db 218 FDTREKYLTLFPHQFI-----SPRTNMTAHTVYVLGV 251  
Qy 190 SGEVFTVSCRVLTRPAPDFEFTLVPTVESKTKPFTLPILTLGELSNSRFPAAIDMLY 249  
Db 252 NRYDQY-----KKHKPMTLVVMVLSPLTVSNTAAPQIKVY 286  
Qy 250 TDPNESIVVQPONGRCTLDTGLTQTLVPTQICAFRGTLISQATARADSTDSPQARNH 309  
Db 287 ANIAPTVV-----HVAGELPSKEGIFPVACADGYGGLVTTDPKTAD----- 327  
Qy 310 PLHVQVKNLDTQY-----DPTDDIPAVLGAIKFGTVFGVASORDVSGQOEGHYA 361  
Db 328 PVYGVKYNPPKTNYPGRFTNLLDVAEACPTFLRFDGKPYVVVTRADDTRLAKFDVSLAA 387  
Qy 362 TRAHEAHIDTTDPKYAPKLGTI---LIKSGDDFNTNQPIRFTVGM 405  
Db 388 KHSNTYLSGIAQYVYQYSGTINLHFMFTGTDSKARYMAYIPPGV 434

RESULT 7  
US-09-738-626-6888  
; Sequence 6888, Application US/09738626  
; Publication No. US20020197605A1  
; GENERAL INFORMATION:  
; APPLICANT: NAKAGAWA, SATOSHI  
; APPLICANT: MIZOGUCHI, HIROSHI  
; APPLICANT: ANDO, SEIKO  
; APPLICANT: HAYASHI, MIKIRO  
; APPLICANT: OCHIAI, KEIKO  
; APPLICANT: YOKOI, HARUHIKO  
; APPLICANT: TATEISHI, NAKO  
; APPLICANT: SENOH, AKIHIRO  
; APPLICANT: IKEDA, MASATO  
; APPLICANT: OKAZAKI, AKIO  
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES  
; FILE REFERENCE: 249-125

US-09-738-626-6888  
; CURRENT APPLICATION NUMBER: US/09/738,626  
; CURRENT FILING DATE: 2000-12-18  
; PRIOR APPLICATION NUMBER: JP 99/377484  
; PRIOR FILING DATE: 1999-12-16  
; PRIOR APPLICATION NUMBER: JP 00/159162  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: JP 00/280988  
; PRIOR FILING DATE: 2000-08-03  
; NUMBER OF SEQ ID NOS: 7059  
; SOFTWARE: Patent in ver. 3.0  
; SEQ ID NO 6888  
; LENGTH: 1344  
; TYPE: PRT  
; ORGANISM: Corynebacterium glutamicum  
US-09-738-626-6888

Query Match 4.2%; Score 121; DB 9; Length 1344;  
Best Local Similarity 23.1%; Pred. No. 0.065;  
Matches 109; Conservative 52; Mismatches 166; Indels 144; Gaps 29;

Qy 23 NDEVMALPEVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSPRNSPGEMLNLELG 82  
Db 188 NEAVWAAP--GDVITQP--GKSLVF--NIK---NGPNDATAADENA--EYGNLEAG 235  
Qy 83 PELNPLYLSHRMYNGVAGQMVOQ-----VVLGNAFTAG-KLIIFAAVPPHPVVISAA 136  
Db 236 KDLVEISS--CGMANGTARGMQIQTNTGHIIVNRGFYNNMAGASDVKANEGHFAVDESLL 293  
Qy 137 QITWCPHVIVDVQLEPVLPLDIRNRPYHYNQENTPRMLVAMLYTLPLR---ANSGED 193  
Db 294 KQTL-----VSGGAP-----TPGTVYTSQIPNPLSAVIADSSVP 327  
Qy 194 VFTVSCRVLTRPAPDFEFTLVPTVESKTKPFTLPIL-TLGELS-----NSR 240  
Db 328 LITDNTATSNPAEPFTFAFNITDDVQVTA--TLHVTSSAGAAATININTEDDGSGFNWA 385  
Qy 241 FPA-----IDMLYTDPNESIVVQPONGRCTLDTGLTQTLVPTQICAFRGTLIS 291  
Db 386 LPAADLTGKSFVFTVTTATDGFNSVTTEPV--RVVDG-----ANTDPLRLNEENQWVS 438  
Qy 292 QTARAADSTDSPQARNHPLHVQVKNLDTGTYDPTDDIPAVLGAIKFGTVFGV-ASQD 350  
Db 439 GTTVDIGASD---VFGDKLELLI-----DDAPAVTNSLSLSAAPTFAMEVTQTD 483  
Qy 351 V-----SQO-----EQHYATRAHEAHIDTTDPKY-----APKL 380  
Db 484 VFFRNGILLAGGEELRIFDQGTI---ANTETISTVPPLVPLHINEDGTLTVSVVAGTKAAPEI 540  
Qy 381 GTILIKSGDDFNTNQPIR-----FTPVGMGDNNRQW-ELPDYSGRL 422  
Db 541 D---LNENDDFOIRN-LRLILPDGRTLTLPAGISDSN--AWLNMGDSAGKL 585

RESULT 8  
US-10-043-487-383  
; Sequence 383, Application US/10043487  
; Publication No. US20030055220A1  
; GENERAL INFORMATION:  
; APPLICANT: HYBRIGENICS  
; APPLICANT: PIERRE, LEGRAIN  
; TITLE OF INVENTION: Protein-protein interactions between Shigella Flexneri polypeptides  
; TITLE OF INVENTION: mammalian polypeptides  
; FILE REFERENCE: B4778A  
; CURRENT APPLICATION NUMBER: US/10/043,487  
; CURRENT FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/261,130  
; PRIOR FILING DATE: 2001-01-12  
; NUMBER OF SEQ ID NOS: 561  
; SOFTWARE: Patent in version 3.1  
; SEQ ID NO 383  
; LENGTH: 2809  
; TYPE: PRT  
; ORGANISM: Shigella Flexneri

US-10-043-487-383

Query Match 4.1%; Score 119; DB 14; Length 2609;  
Best Local Similarity 21.0%; Pred. No. 0.3;  
Matches 131; Conservative 64; Mismatches 219; Indels 209; Gaps 31;

QY 48 IDPWIREN-----FVQAPQ-GEFTVSPR-----NSPGEMLLNLELGPBLNPLYL--S 90  
DB 1658 VDDVVENEDGTDFIYTAPOPGKTVICVRFGEHVPNSPFQVOTALAGDQPSVQPPLSRQ 1717  
QY 91 HLSRMVNGYAGGMQVQVVLGNAGNATAGKIIFAAVPPHPFVENISAAQITMCPHVIVDVQR 150  
DB 1718 QLAPOYTAQGGQQTWA-----PERPLVGNG-----LDVTS 1749  
QY 151 LEPVLLPLDIRNRFHYNQENTPMRL-----VAMLYTFLRANSGE- 192  
DB 1750 LRPFDLVIP-----FTIKKEITGEVRMPGSKVAQPTITDNKDGTVTVRYAPSEAGLHEM 1804  
QY 193 -----DVFTVSCRVLTRPADPFEFTFLVPPVTSKTKPFTLPILTLGL 236  
DB 1805 DIRYDNMHIPGSPLOFYDYVNCGHVTAYGPG-----LTHGVNKNPATFTVNTKDAGEG 1858  
QY 237 SNS-----RPPAIDLMLYTDPE-----SIWVQ-----PQNGRCTLDG 269  
DB 1859 GLSLAIEGSKAEISCTDNQDGTCSVSLPVLPGDYSILVKYNEQHVPGSPFTARVTGDD 1918  
QY 270 TLQ-----GTTQLVPTQICAPRGTLISQARAADSTDS-----QRARN----- 308  
DB 1919 SMRSHLKVGSAAIPINISETDLSTLTATVVPSPGREPCLLKRLRNGHVGISFVPKET 1978  
QY 309 --HPLHVQKNLDGTQYDPTDDIPAVLGAIDPKGTGFGVASORDVSGQ-QEKGHYATRAH 365  
DB 1979 GEHLVHVK-KN---GQHVASSPFPVVISQSE-----IGDASRVVSGQGLHEGHTFEPA- 2028  
QY 366 EAHIDTTPKAPKLTGILIKSGSD--DENTNQPIRFTPVGMGDNMNMWQWELPDYSGLRT 423  
DB 2029 EFIDTRDAGY---GLSLIEGSKVDINTED-----LEDGTCRVYCTPBGNYI 2077  
QY 424 LNMNLAPAVSPFP-----GE-----RILFFRSIVPSAGGYSGY-IDCLIPQEWVQHF- 471  
DB 2078 INIKFADQHVPGSPSVKVTGEGRVKESITRRRRAPSVANVGSCHDLSLKPEISIQDWT 2137  
QY 472 YOEAAAPS-----QSAVALRVYVNPDTGRNIFEAKLHREGFLTVANGCNPPIVVP 520  
DB 2138 AQVTSFGSKTHEAEIVEGENHTYCIKRFVPAEMGTHTVSVKYGQHV-----2183  
QY 521 PNGYFRFEAGNQFYTLAPMGSG 543  
DB 2184 PGSPFQF-----TVGPLGEG 2198

RESULT 9

US-10-231-956A-87  
; Sequence 87, Application US/10231956A  
; Publication No. US2004005323A1  
; GENERAL INFORMATION:  
; APPLICANT: Lorens, James B.  
; APPLICANT: Xu, Weiduan  
; APPLICANT: Bogenberger, Jakob  
; APPLICANT: Holland, Sacha  
; APPLICANT: Rigel Pharmaceuticals, Incorporated  
; TITLE OF INVENTION: Modulators of Angiogenesis  
; FILE REFERENCE: 021044-004100US  
; CURRENT APPLICATION NUMBER: US/10/231,956A  
; CURRENT FILING DATE: 2001-08-30  
; NUMBER OF SEQ ID NOS: 522  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 87  
; LENGTH: 2647  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-231-956A-87

Query Match 4.1%; Score 119; DB 12; Length 2647;  
Best Local Similarity 21.0%; Pred. No. 0.3;  
Matches 131; Conservative 64; Mismatches 219; Indels 209; Gaps 31;

QY 48 IDPWIREN-----FVQAPQ-GEFTVSPR-----NSPGEMLLNLELGPBLNPLYL--S 90  
DB 1696 VDDVVENEDGTDFIYTAPOPGKTVICVRFGEHVPNSPFQVOTALAGDQPSVQPPLSRQ 1755  
QY 91 HLSRMVNGYAGGMQVQVVLGNAGNATAGKIIFAAVPPHPFVENISAAQITMCPHVIVDVQR 150  
DB 1756 QLAPOYTAQGGQQTWA-----PERPLVGNG-----LDVTS 1787  
QY 151 LEPVLLPLDIRNRFHYNQENTPMRL-----VAMLYTFLRANSGE- 192  
DB 1788 LRPFDLVIP-----FTIKKEITGEVRMPGSKVAQPTITDNKDGTVTVRYAPSEAGLHEM 1842  
QY 193 -----DVFTVSCRVLTRPADPFEFTFLVPPVTSKTKPFTLPILTLGL 236  
DB 1843 DIRYDNMHIPGSPLOFYDYVNCGHVTAYGPG-----LTHGVNKNPATFTVNTKDAGEG 1896  
QY 237 SNS-----RPPAIDLMLYTDPE-----SIWVQ-----PQNGRCTLDG 269  
DB 1897 GLSLAIEGSKAEISCTDNQDGTCSVSLPVLPGDYSILVKYNEQHVPGSPFTARVTGDD 1956  
QY 270 TLQ-----GTTQLVPTQICAPRGTLISQARAADSTDS-----QRARN----- 308  
DB 1957 SMRSHLKVGSAAIPINISETDLSTLTATVVPSPGREPCLLKRLRNGHVGISFVPKET 2016  
QY 309 --HPLHVQKNLDGTQYDPTDDIPAVLGAIDPKGTGFGVASORDVSGQ-QEKGHYATRAH 365  
DB 2017 GEHLVHVK-KN---GQHVASSPFPVVISQSE-----IGDASRVVSGQGLHEGHTFEPA- 2066  
QY 366 EAHIDTTPKAPKLTGILIKSGSD--DENTNQPIRFTPVGMGDNMNMWQWELPDYSGLRT 423  
DB 2067 EFIDTRDAGY---GLSLIEGSKVDINTED-----LEDGTCRVYCTPBGNYI 2115  
QY 424 LNMNLAPAVSPFP-----GE-----RILFFRSIVPSAGGYSGY-IDCLIPQEWVQHF- 471  
DB 2116 INIKFADQHVPGSPSVKVTGEGRVKESITRRRRAPSVANVGSCHDLSLKPEISIQDWT 2175  
QY 472 YOEAAAPS-----QSAVALRVYVNPDTGRNIFEAKLHREGFLTVANGCNPPIVVP 520  
DB 2176 AQVTSFGSKTHEAEIVEGENHTYCIKRFVPAEMGTHTVSVKYGQHV-----2221  
QY 521 PNGYFRFEAGNQFYTLAPMGSG 543  
DB 2222 PGSPFQF-----TVGPLGEG 2236

RESULT 10

US-10-677-558-2  
; Sequence 2, Application US/10677558  
; Publication No. US20040096972A1  
; GENERAL INFORMATION:  
; APPLICANT: AUDIT, Muriel  
; APPLICANT: COSET, Francois-Loic  
; TITLE OF INVENTION: CHIMERIC PLASMID COMPRISING A REPLICATIVE RETROVIRAL GENOME AND  
; FILE REFERENCE: 1759.135  
; CURRENT APPLICATION NUMBER: US/10/677,558  
; CURRENT FILING DATE: 2003-10-02  
; PRIOR APPLICATION NUMBER: PCT/FR02/03934  
; PRIOR FILING DATE: 2002-11-18  
; PRIOR APPLICATION NUMBER: FR 0114976  
; PRIOR FILING DATE: 2001-11-20  
; NUMBER OF SEQ ID NOS: 15  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 1199  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: product of pol gene of pAM plasmid  
US-10-677-558-2

Db	30	QLRYSVVESEPGTLVGNVAQDLGLKMTDLLSRRLQSGSENGRYFSLSLMSGALAVNQK	89
Qy	105	VQVVLGNAFTAGKIIFAAVPPHPPVENISAAQTMCPHVIVDVROLEPVLPLPDIRNR	164
Db	90	IDRESLCGASTSCLLPVQVWTEH-PLELI-----RVEVELDLDNDSFS	132
Qy	165	PFHYNQENTPMRLVAMLYTPLRANGSDVFTVSCRVLTRPAPDFEFTFLVPPTVEKTK	224
Db	133	FATPEREMRISSAAGARFPLDSAQDPDVGNTVSVFYL-SPNSHFSLVNKTILKOG--K	189
Qy	225	PFTLPITILGELSNRFPAAIDMLYDNE-----SIVVQONGRCTLDGTLQG	273
Db	190	PF--PELVLEQLDREAQARHQLVLTAVDGGTPEARSGTTLISVILVDINDNAP---TFQS	244
Qy	274	TTQLVPTQICAFRGTLISQTAABAOSTDQSPQARNHPL-----HVQ--VKNLDGTQYDP	325
Db	245	SVLRVGPENAPITGTLRLR-----LNATD-PDEGTNGQLDYSFGDHTSEAVRNLFQ--LDP	297
Qy	326	TDDIPAVLGAIDFKGTVFGVASQDVSGQEQGHYATRAHEAHIDTTPKYPAPKLTILI	385
Db	298	SSGAIHVLGPIDPE-----ESRFEIHAHARDQGGQPMGHCVI	336
Qy	386	KSGSDDFNTNQ-----IRETPVG--MGDNWQWELPDYSGRLTLNMLAPAV	432
Db	337	QVDVGDVNDNAPEVLLASLANPVLESTPGTVGVGLFNVDRD-----SGR--NGEVSLDI	389
Qy	433	SPSPGGERILFFRSIVPSAGGY-----SGYIDCLIP-----	464
Db	390	SPDLFPQ-----IKPENHYSLTQPLDREATSHYIIELLASDAGSFLKHLITRL	442
Qy	465	-----QEWQHFYQEAAPSQSAVALRVYNPDGTGRNIPEAKLHREGFLTIVAN	511
Db	443	NISVDNDNAPRFNQQLYTAVILENRPGPSLLCTVAASDDPTGDN-----ARLTYSI	493
Qy	512	CNNPIVVPPNGY	524
Db	494	VGNQVQGAPASSF	506
RESULT 12			
US-10-156-761-7936			
; Sequence 7936, Application US/10156761			
; Publication No. US20030119018A1			
; GENERAL INFORMATION:			
; APPLICANT: OMURA, SATOSHI			
; APPLICANT: IKEDA, HARUO			
; APPLICANT: ISHIKAWA, JUN			
; APPLICANT: HORIKAWA, HIROSHI			
; APPLICANT: SHIBA, TADAYOSHI			
; APPLICANT: SAKAKI, YOSHIYUKI			
; APPLICANT: HATTORI, MASAHIRO			
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES			
; FILE REFERENCE: 249-262			
; CURRENT APPLICATION NUMBER: US/10/156,761			
; CURRENT FILING DATE: 2002-05-29			
; PRIOR APPLICATION NUMBER: JP 2001-204089			
; PRIOR FILING DATE: 2001-05-30			
; PRIOR APPLICATION NUMBER: JP 2001-272697			
; PRIOR FILING DATE: 2001-08-02			
; NUMBER OF SEQ ID NOS: 15109			
; SEQ ID NO 7936			
; LENGTH: 368			
; TYPE: PRT			
; ORGANISM: Streptomyces avermitilis			
US-10-156-761-7936			
Query Match 3.6%; Score 104; DB 14; Length 368;			
Best Local Similarity 23.2%; Pred. No. 0.38;			
Matches 82; Conservative 39; Mismatches 120; Indels 112; Gaps 18;			
Qy	16	ANLVPEANDEWMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAPGEGFTVSPRNSPGEM	75

Db	24	DEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAPGEGFTVSPRNSPGEMLINLEGP	83
Db	459	DOQKAYQEIQKALLTAPALGULPDLTKPF--ELFVDEKQGY-----ANGVLTKLGP	507
Qy	84	ELNPYLSHLSRMVYAGG-----MQVQVVLGNAFTAGKIIFAAVPPHPPVENISAAQIT	139
Db	508	WRPP--VAYLSKKLDVPAAGWPPCLRWAAIAVLTKDAGKL-----TWGQPLVI	554
Qy	140	MCPHVIVDVROLEPVLPLPD--IN-REFHYNQENTPMRLVAMLYTPLRANGSDVFT	196
Db	555	LAPHAV-----EALVKQPPDRWLSNARWTHQ-----ALLDTRVQFGP---	594
Qy	197	VSCRVLTRPAPDFEFTFLVPPTVESKTPFTLPIL-----TLGELSNRFPAAIDMLYTD	251
Db	595	---VVALNPA-----TLL--PLPEEGLQHDCLDILAEAGHTRSDLTDOPLPDADHTWYTD	644
Qy	252	PNBSIVVQONGRCTLDGTLGTTQLV-----PTQICAFRGTLI--SOTARAADS-----	299
Db	645	GSSFL-----QEQORKAGAAVTTETEVIMARALPAGTSAORAEILIALTQALKWAECKLVN	700
Qy	300	TDSPQARNHPLHVQV-----KNLDGTQYDPTDDIPAVLGAI-----	336
Db	701	YDSRVAFATAHGEIYRERGLTSEGKEIKNKDEILALLKALFLPKRLSIHCPGHQK	760
Qy	337	----DFKGTVFGVASQDVSGQEQ-----HYATRAHE-AHIDTTPKYAPKLTIL	384
Db	761	GNSAARGNRMAQAAAREVATRETPTGTSLLIENSTPYTHEHPTVTTDKLTKLGATY	820
Qy	385	IKSGSDDFNTNQI---RFTPGVGMGNWQWELPDYSGRLT-LNNMLAPAV---SPS---	435
Db	821	DSAKKYVYQKVPNDQFT-----FELDLFLHQLTHLSFSKTKALLERSPSPY	869
Qy	436	-----PFGERI-----LFFRSIVPSAGGYGS	456
Db	870	YMLNRDRLKNIETCKACQVNASKSAVKQGTVRGHRFGTWHWEIDFTEVKP--GLYGY	927
Qy	457	GYIDCLIP--QEWQHFYQEAAPSQSAVALRVYNPDGTGRNIPEAKLHREGFLTIVANCGN	514
Db	928	KYLLVVDVTFSGWI-----EAPFKKETAKV-----VTKKLLSEIIPFRGMPQVLGTDN	976
Qy	515	NPVV	519
Db	977	GPAFV	981
RESULT 11			
US-10-174-677-101			
; Sequence 101, Application US/10174677			
; Publication No. US20030190704A1			
; GENERAL INFORMATION:			
; APPLICANT: Xie, Ting			
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR ANCHORING STEM CELLS IN A MICROENVIR			
; FILE REFERENCE: 40716(IP-012)			
; CURRENT APPLICATION NUMBER: US/10/174,677			
; CURRENT FILING DATE: 2002-06-19			
; NUMBER OF SEQ ID NOS: 117			
; SOFTWARE: Patentin version 3.1			
; SEQ ID NO 101			
; LENGTH: 944			
; TYPE: PRT			
; ORGANISM: Homo sapiens			
US-10-174-677-101			
Query Match 3.6%; Score 104.5; DB 14; Length 944;			
Best Local Similarity 21.5%; Pred. No. 1.6;			
Matches 119; Conservative 56; Mismatches 213; Indels 165; Gaps 22;			
Qy	61	QGEFTVSPRNSPGEMLIN-----LELGFELNPLYLSHLSRMVYAGGMQ	104



Db 56 AELLPOG-----ALDPPARAQIVLTVGRSTVMDGVTRHEVPSAE-----PQQA 102  
Qy 76 LLNLELGPENLNPYLHLSRM-----YNGYAGGMQVQVVLVAGNAFTAGKIIFAAVP 125  
Db 103 KLTLL-TGED-----VSRMDVDFSGVKYFGLPPEERVAVIVA--KYTAGLVPAIVP 152  
Qy 126 PHF-----PVENISAQITMCPHIVDVROLEP-----VLLPLDIRRFFHYNGEN 172  
Db 153 SVFMVAPPLQAVPSHQGTDLAY-----VRQLAAQAGYQFFVAPGPLGV--NIAHWPEN 206  
Qy 173 TPRLVAMLYTPLRANGSE-----DVFTVSCRVLTRPADPFEFTLVP-PTVESKTKP 225  
Db 207 KGLPQELLVNCDASNVESLSFDFGSATQYVVLQNPKFPVPVPVDPVSPPLSP 266  
Qy 226 FTLPLITLGLSNSRFPAAIDMLYTDPNESIVVQPNCRCTLDGTLGGTTLQVLPQICAF 285  
Db 267 M-----GRRRLPLK-----VQS-----IHGLAKLTPQAAGV 294  
Qy 286 RGLTISQARAADSTD-----SPQARNH-PLHVQVKNLDTQY 323  
Db 295 -----ALARAADSFEVSGSLDVLRYGQPLRARPPVPHGAGAAYDGTYY 341

## RESULT 13

US-10-104-047-3560  
; Sequence 3560, Application US/10104047  
; Publication No. US20030236392A1  
; GENERAL INFORMATION:  
; APPLICANT: HELIX RESEARCH INSTITUTE  
; TITLE OF INVENTION: No. US20030236392A1el full length cDNA  
; FILE REFERENCE: H1-A0105  
; CURRENT APPLICATION NUMBER: US/10/104,047  
; CURRENT FILING DATE: 2002-03-25  
; PRIOR APPLICATION NUMBER:  
; PRIOR FILING DATE:  
; NUMBER OF SEQ ID NOS: 4096  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 3560  
; LENGTH: 394  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-104-047-3560

Query Match 3.6%; Score 103.5; DB 15; Length 394;

Best Local Similarity 22.7%; Pred. No. 0.48;  
Matches 49; Conservative 32; Mismatches 66; Indels 69; Gaps 9;

Qy 129 PVENISAQITMCPHIVDVROLEPVLLPL-----DIRRFFHYNGENTPRMLVAML 182  
Db 73 PNPVSCATPSLLSLVLKCPCLSLIPLPGRKCTYGIKCRFFHPRPSCPQ-RSVA-- 129  
Qy 183 YTPLRANGSDVFTVSCRVLTRPADPFEFTLVPPTVESKTKPFTPLITLGLSNSRPP 242  
Db 130 -DELRA-----LLSPRPSKDK-----NGRRP 153  
Qy 243 AAIDMLYTDPNESIVVQPNCRCTLDGTLGGTTLQVLPQICAFRGTLLISQARAADSTD 302  
Db 154 -----SPSSQSSLLTSEQSLDGKLG-AQASP-----GSRQGLTQTVAPSGRSLA 201  
Qy 303 PQARNHPLHVQVKNLDTQYDPTDDIPAVLGAIDF 338  
Db 202 PSGG-----SGSSFGPTDLPQLTDLSPY 225

## RESULT 14

US-10-080-334-200  
; Sequence 200, Application US/10080334  
; Publication No. US20040002584A1  
; GENERAL INFORMATION:  
; APPLICANT: Pena, Carol E. A.  
; APPLICANT: Shimkets, Richard A  
; APPLICANT: Li, Li  
; APPLICANT: Shenoy, Suresh G

; APPLICANT: Kekuda, Rameesh  
; APPLICANT: Spytek, Kimberly A.  
; APPLICANT: Vernet, Corine A. M.  
; APPLICANT: Malyankar, Uriel M.  
; APPLICANT: Guo, Xiaojia  
; APPLICANT: Gusev, Vladimir Y  
; APPLICANT: Casman, Stacie J  
; APPLICANT: Boldog, Ferenc L  
; APPLICANT: Furtak, Katarzyna  
; APPLICANT: Tchernev, Velizar T  
; APPLICANT: Patturajan, Meera  
; APPLICANT: Gangolli, Esha A  
; APPLICANT: Padigaru, Muralidhara  
; APPLICANT: Liu, Xiaohong  
; APPLICANT: Baumgartner, Jason C.  
; APPLICANT: Gerlach, Valerie  
; APPLICANT: Spaderna, Steven K  
; APPLICANT: Zerhusen, Bryan D  
; TITLE OF INVENTION: Proteins, Polynucleotides Encoding Them and Methods of  
; FILE REFERENCE: 21402-275  
; CURRENT APPLICATION NUMBER: US/10/080,334  
; CURRENT FILING DATE: 2002-02-21  
; PRIOR APPLICATION NUMBER: 60/270,523  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: 60/322,712  
; PRIOR FILING DATE: 2001-09-17  
; PRIOR APPLICATION NUMBER: 60/311,980  
; PRIOR FILING DATE: 2001-08-13  
; PRIOR APPLICATION NUMBER: 60/330,307  
; PRIOR FILING DATE: 2001-10-18  
; PRIOR APPLICATION NUMBER: 60/278,796  
; PRIOR FILING DATE: 2001-03-26  
; PRIOR APPLICATION NUMBER: 60/281,521  
; PRIOR FILING DATE: 2001-04-04  
; PRIOR APPLICATION NUMBER: 60/276,677  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/311,595  
; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: 60/270,220  
; PRIOR FILING DATE: 2001-02-21  
; PRIOR APPLICATION NUMBER: 60/274,295  
; PRIOR FILING DATE: 2001-03-08  
; PRIOR APPLICATION NUMBER: 60/291,765  
; PRIOR FILING DATE: 2001-05-17  
; PRIOR APPLICATION NUMBER: 60/270,797  
; PRIOR FILING DATE: 2001-02-23  
; PRIOR APPLICATION NUMBER: 60/276,400  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/270,810  
; PRIOR FILING DATE: 2001-02-23  
; NUMBER OF SEQ ID NOS: 388  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 200  
; LENGTH: 3317  
; TYPE: PRT  
; ORGANISM: Rattus norvegicus  
US-10-080-334-200

Query Match 3.5%; Score 103; DB 15; Length 3317;

Best Local Similarity 19.2%; Pred. No. 17;  
Matches 136; Conservative 73; Mismatches 242; Indels 256; Gaps 34;

Qy 14 GAANLVPEANDEVMALEPVVGASIAAPVVGQNIIDPWIRENFVQAP---QGSEFTVSPRN 70  
Db 1703 GRYTLIVTATDQC----FILSHRLTSTTTVLNVND--INDNVPTFPRDYGEPDVT-BG 1755  
Qy 71 SPGEMLLNLELGPENLNPYLHLSRMVNGYAGGMQVQVVLVAGNAFTAGKIIFAAVPDPFPV 130

1756 QP-----GPRVWTFPLAH--DRDSCPNQGVSVV---DGDPLGEFVISPVEGLRV 1801  
131 -----ENISAAQITMCPH-----VIVDVQLEFVLLPLPDIRNFFH 167  
1802 KDVLDRETIATFNLTICARDGVPPLSSMLVGIRVLDINDNDPVLNLP-----1853  
168 YN--QENTPRMVLY-----AMLYTPLRANSQDVF-----TVSCRVLTRP-- 205  
1854 MNITISENSPVSSFVAHVLAADSGCNALTFNITAGNRBARAFINATTGIIVTNPLD 1913  
206 ---APDEFETFLV---PPTVESKTPFTPLITLIGELSNRFPAAIDMLYTDPNESIVVQ 259  
1914 RERIEFRLTVSKONPENPRIARKDFDLVSLAD-ENDNHP-----LFTEGTYQAEVM 1967  
260 PONGRCTLDTLQ-----TTOLVPTQICAF-----RGTLLSQ 292  
1968 ENSPACTPLTVLNGPILALDADEAVVVTYQLLGHSHSLFVIDNSTGVTVRSVILDR 2027  
293 TABAADSTSPQARNHPLHVQKNLDGTQY-----DPTDDIPAV-----332  
2028 EA-----FSPFLELLLAEDVGQLNGTAYLFTITLDDNDNWPFSPPAYTVHLENC 2081  
333 ---LGAIDFKGTGVFVASORDVSGQEQG--HYAT---RAHEAHID-----370  
2082 PGFSVLQITATDSDGLNGELVYRIEAGAQDRFLHPVTVGVRVGNATIDREEQESYRLT 2141  
371 --TTPDKYAPKLTILIKSGSDDFNTNP-----IRFTPVGMGD 407  
2142 VVATDRGTVPLSGTATVITLIDINDSRPEFNPIQTVSVLESTEGTVIANVTAIDL 2201  
408 NNRQWEL-----PQYSGRLTLMNLA-----429  
2202 NPKLEYHLSIVAKDDTRLVDPQEDAFVNTINTGVSIVKSNRELNAVTEYVTLNVION 2261  
430 ---PAVSPSPGGR-----ILFRSIVPSAGGSGYIDCLIPQEWVQHFYQAAAPSQA 481  
2262 ASDLPERSVSPNAKLTWNILVDNDNTPOKPGIITY-----YTERVLEGATPGTTL 2313  
482 VALVRYNPDTRGRNIFEAKLHREGFLTVANGCNPPIVVPNGYFRFE 528  
2314 IA-VNAVDPDKGLN-----GLITVTLDD-----LIPP-GYVQLE 2345

RESULT 15  
US-10-282-122A-62618  
; Sequence 62618, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA 034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347

PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625  
; PRIOR FILING DATE: 2000-11-27  
; PRIOR APPLICATION NUMBER: 60/257,931  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: 60/267,636  
; PRIOR FILING DATE: 2001-02-09  
; PRIOR APPLICATION NUMBER: 60/269,308  
; PRIOR FILING DATE: 2001-02-16  
; Remaining prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 78614  
; SOFTWARE: Patent version 3.1  
; SEQ ID NO 62618  
; LENGTH: 1167  
; TYPE: PRT  
; ORGANISM: Mycobacterium bovis  
US-10-282-122A-62618

Query Match 3.5%; Score 102; DB 12; Length 1167;  
Best Local Similarity 19.3%; Pred. No. 3.9;  
Matches 105; Conservative 73; Mismatches 211; Indels 154; Gaps 26;

QY 19 VPEANDEWMALEPVVGASTAAFPVVGQONIIDPWIRENFVQAPQGEF-----TVSPR 69  
DB 310 VPDGHDQASWL-----RHEVDAGLRRRFPAGPDGGRERAAYEIDVICSK 354  
QY 70 NSPEMLLNLELGPENLPYLSHLMYVNGYAGGMQVVLVGNAGNAFTAGKIIFAAVPPHP 129  
DB 355 GFPSYFLLIVADL-----ISVARSAGIRVG-PGRGSAAGSLVAYA-----392  
QY 130 VENISAAQITMCPHVIYDVQLEFVLLPLPDI-----RNRFFHYNQENTPRMLV-- 179  
DB 393 ---LGITDIDIPHGLLPERFLNPERTSPMDIDIDFDRRRGEMVRYAADKMGHDRVAVQ 449  
QY 180 -----AMLYTPLRANSQDVFVTSRVLTRPADPDEFETFLVPPTVESKTKPFTLP 230  
DB 450 ITFGTIKTAALKDSARIHYGQGFADRI-----TKALPPAIMAK-----DIP 495  
QY 231 LTLGELSNRFPAAID---MLYTDPNESIVVQPNQGRCTLDGTLOGTTLVPTQICAFRG 287  
DB 496 SGITDPSHERYKEAAEVRGLIETDPVRTIYQTARG-----LEGLRNAGVHACA--- 545  
QY 288 TLISOTARAADSTDSQPARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIDFKG-----TVF 343  
DB 546 -VIMSSEPLTEAIPLMKRPQDGA-----ITGWDYPACEAIG--LLKMDPLGLRNLTI 596  
QY 344 G-----VASORDVSGQEQGHYATRAHEAHIDTTDPKYAPKLTILIKSG--SDDFNTNQ 396  
DB 597 GDAINDVRANRGIDLDLESVPLDDKATVELLGRGD-----TLGVFQLDGGPMRDLRRMQ 651  
QY 397 PIRFT-----PVGMDNNRWQWELPDYSGRLTLNMLAPAVSPGPFGERILFF 444  
DB 652 PTGFEDVAVAVIALYRFGPMGMAHN-----DYADR-----KNNFQAIKPH-ELERPL 699  
QY 445 RSIVPSAGYGVGYIDCLIPQEWVQHFYQAAAPSQAVA-LVRYNPDTRGRNIFEAKLHR 503  
DB 700 REIL--AETVG-----LIVYQIMRIQKVASYSLARADILRKAMGKKKREVLEKEF-- 750  
QY 504 EGF 506  
DB 751 EGF 753

Search completed: June 1, 2004, 14:04:46  
Job time : 37.1594 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.6443 Seconds  
(without alignments)  
4310.252 Million cell updates/sec

Title: US-09-926-799-10  
Perfect score: 2906  
Sequence: 1 MKMASNDAPNSDGAANLVP.....GNQFYTLAPMSGGQRRRAQ 550

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1830.5	63.0	539	1 COAT LORDV	P54635 lordedale v
2	1089.5	37.5	546	1 COAT SOUV3	Q04542 southampton
3	293	10.1	702	1 COAT SMSV1	P36284 san miguel
4	290	10.0	2344	1 POLN RHDV	P27410 rabbit hemo
5	287.5	9.9	703	1 COAT SMSV4	P36285 san miguel
6	283	9.7	668	1 COAT FCVF4	P27405 feline cali
7	280.5	9.7	671	1 COAT FCVF9	P27406 feline cali
8	270.5	9.3	668	1 COAT FCVC6	P27404 feline cali
9	236.5	8.1	2208	1 POLN MANCV	Q69014 manchester
10	129.5	4.5	2332	1 POLG FMDVO	P03305 f genome po
11	125.5	4.3	2194	1 POLG HE701	P32537 h genome po
12	124.5	4.3	2333	1 POLG FMDV1	P03306 f genome po
13	119	4.1	2647	1 FLNA_HUMAN	P21333 homo sapien
14	118.5	4.1	2193	1 POLG CX16T	Q9qf31 c genome po
15	118	4.1	2336	1 POLG FMDV2	P49303 f genome po
16	117.5	4.0	3375	1 UNS2 CABEL	Q06561 caenorhabdi
17	117	4.0	4289	1 TENX_HUMAN	P22105 homo sapien
18	116	4.0	2193	1 POLG CX16G	Q65900 c genome po
19	115	4.0	1034	1 ADD DROME	P54362 drosophila
20	114.5	3.9	788	1 BCBS_XANAC	P58933 xanthomonas
21	114	3.9	2194	1 POLG EC30B	Q9wr78 e genome po
22	114	3.9	2647	1 FLNA_MOUSE	Q8btm8 mus musculus
23	113	3.9	855	1 POLG HRV3	Q82081 human rhino
24	111.5	3.8	2175	1 POLG BOVEV	P12915 b genome po
25	111.5	3.8	2206	1 POLG_POL32	P06209 poliovirus
26	110.5	3.8	1204	1 POL MLVFS	P26810 friend muri
27	110.5	3.8	1204	1 POL MLVFF	P26809 friend muri
28	110.5	3.8	2332	1 POLG FMDVA	P03308 f genome po
29	109.5	3.8	1199	1 POL MLVMO	P03355 moloney mur
30	108.5	3.7	798	1 PBPA_NEICI	Q86088 neisseria c
31	108.5	3.7	1204	1 POL MLVFP	P26808 friend muri
32	107	3.7	1011	1 POLG FMDVT	P15072 foot-and-mo
33	107	3.7	2201	1 POLG_CXA9	P21404 c genome po

RESULT 1									
COAT LORDV	COAT LORDV	STANDARD;	PRT;	539 AA.					
AC	P54635;								
DT	01-OCT-1996	(Rel. 34, Created)							
DT	01-OCT-1996	(Rel. 34, Last sequence update)							
DT	01-NOV-1997	(Rel. 35, Last annotation update)							
DE	Coat protein (Capsid protein).								
OS	Lordedale virus (Human enteric calicivirus).								
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;								
OC	Norovirus.								
OX	NCBI_TaxID=82658;								
RN	[1]	SEQUENCE FROM N.A.							
RP	MEDLINE=96005060; PubMed=7561776;								
RA	Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;								
RT	"Human enteric Caliciviridae: the complete genome sequence and								
RT	expression of virus-like particles from a genetic group II small								
RT	round structured virus."								
RL	Genbank accession: F652395.2355(1995).								
CC	- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See http://www.isb-sib.ch/announce/								
CC	or send an email to license@sib-sib.ch).								
CC	-----								
DR	EMBL; X86557; CAA60255.1;								
DR	InterPro; IPR004005; Calici_coat.								
DR	InterPro; IPR008975; Viral_Cap_coat.								
DR	Pfam; PF00915; Calici_coat; 1.								
KW	Coat protein; Glycoprotein.								
FT	CARBOHYD 172 172								
FT	CARBOHYD 307 307								
FT	CARBOHYD 406 406								
FT	CARBOHYD 531 531								
SQ	SEQUENCE 539 AA; 58775 MW; 5E5C6E7F2C5PD21 CRC64;								
Query Match 63.0%; Score 1830.5; DB 1; Length 539;									
Best Local Similarity 62.11%; Pred.No. 2.3e-130;									
Matches 346; Conservative 83; Mismatches 101; Indels 27; Gaps 6;									
QY	1	MKMASNDAPNSDGAANLVP	EAANLVPEANDEVALEPVVGVGQNIIDPWIRENFVQAP	60					
Db	1	MKMASNDAPNSDGAANLVP	EAANLVPEANDEVALEPVVGVGQNIIDPWIRENFVQAP	60					
QY	61	QGEFTVSPRNSPGEMLLNLELGPENLYSLHLSRMVYAGQVQVVLGNAGFTAGKI	120						
Db	61	QGEFTVSPRNPAPCEILWSAPGLDNLNLYSLHLSRMVYAGQVQVVLGNAGFTAGKI	120						
QY	121	FAAVPHPPFVENISAQITMCPHVIVDVRQLEPVLPLPDINRFFHYNQETPRLVA	180						
Db	121	FAAVPPNFPFTGLSPQVTMFPHIIVRQLEPVLPLPDINRFFHYNQETPRLVA	180						

Matches 250; Conservative 75; Mismatches 173; Indels 89; Gaps 107

1 MKWASNDAAPSNDGAA---NLVPEAND-EWMALEPVVGASIAAPVVGQOQNIIDPWIRENF 56  
 1 MMWASDKAPOSADGAGAGQLVFEVNTADPLPMEPVAGPTTAVATAGQVNMIDPMVNNP 60

57 VQAPQGEFTVSPRNSPGEMLNLLELGPENLYSLHSMRYNGYAGQVQVVLGNAFTA 116  
 61 VQSPQGEFTISPNNTPGDILFDLQGLPHLNPFLSHLSQMYNGVGNRVRILLAGNAFSA 120

117 GKIIIPAAVPPHPFVENISAAQITCMCHVIVDVRLQLEPVLPLPLDIRRPFHYNQENTPRM 176  
 121 GKIIIVCCVPFGFTSSSIIIAQLTFPHVIADVLTLEPIEMLEPDEVNVLVHTK-DNQPTM 179

177 RLVAMLYTLPLRANGS---EDVFTVSCRVLTRPAPDPFEFTLVLPPTVESKTKPTLPLTL 233  
 180 RLVCMLYTLPLATGGSGNSDSFVAGRVLTAPSSDFSFLFLVPTTIEQKTRAFVVPNPL 239

234 GELSNRSPALDMLYTPDNESIIVVQFONGRCTLDGTLOGTQVLVPTQICAFRGTLISQT 293  
 240 QTLNSRSPFLSQMLSPDASQVVFQNGRCLLDGQLGTTATSPATSGQLFVRGK-INQG 298

294 ARAADSTDSQPARNRHLHVQVKNLDTGYDPTDDIPAVLGAIDF-----KG 340  
 299 ARTNLUTE-----VDGRPFMAFDS-PAPVGFPPFGCKDWHMRISKTPNN 341

341 TVFG-----VASQDVSGQEQGHVATRAHEAHIDTTDPKYAPKLGTLIKSGSDDFN-- 393  
 342 TSGDDPMRSVSQTNVQG-----FVPHLGSIQF--DEVFNHP 376

394 -----TNQPIRFTPVGMDNNRWQELPDYSGRLTLNMNLAPAVSPSPGGERILF 443  
 377 TGDYIGTIEWISQP--STPPGTDIN---LWEIPDTGSSLSQANLAPVPFPGFGEALV 431

444 FRISVPSAGYGS-GYIDCLIPQEWVQHFOYQBAAPSQSAVALRVVNPDTGNIEAKLH 502  
 432 FVGAFFQPNRSAPNDVPCLLPQEIYTHFVSEQPTMGDAALLHTVDPDTNRNLGEFKDY 491

503 REGFLTVAA--NCGNNPIVPPNGYCFPEANGNQFVTLAPMGSGQGR 547  
 492 PGGYLTCVPNGVGAGPQLPLNGVFLFVSWSRFYQLKPVGTASTAR 538

RESULT 3  
 COAT SMSV1 STANDARD; PRT; 702 AA.  
 ID COAT SMSV1  
 P36284;  
 01-JUN-1994 (Rel. 29, Created)  
 01-JUN-1994 (Rel. 29, Last sequence update)  
 15-DEC-1998 (Rel. 37, Last annotation update)  
 Coar protein (Capsid protein).  
 DE San Miguel sea lion virus (serotype 1) (SMSV 1).  
 OS Viruses; sRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OX NCBI\_TaxID:36406;  
 [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=92410750; PubMed=1529644;  
 RX Neill J.D.;  
 RA "Nucleotide sequence of the capsid protein gene of two serotypes of  
 RT San Miguel sea lion virus: identification of conserved and non-  
 RT conserved amino acid sequences among calicivirus capsid proteins.";  
 RL Virus Res. 24:211-222(1992).  
 CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALCIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC

Mat

QY	181	MLYTPLEA-NSGSDVFTVSCRVLTRAPDPFTFLVPPTVESKTKPFTLPILTLGELSNS	239
DB	181	MLYTPLEA-NSGSDVFTVSCRVLTRAPDPFTFLVPPTVESKTKPFTLPILTLGELSNS	240
QY	240	RFPNADMLTVDNESIVQVQNGRCTLDTGLTQGTTLQVPTQICAFRGTLSIQARAADS	299
DB	241	RFPNADMLTVDNESIVQVQNGRCTLDTGLTQGTTLQVPTQICAFRGTLSIQARAADS	288
QY	300	TDSQPARNHLHVQVKNLQDGTQYDPTDDIPAVLGAIDFKGTGFGVQASQDVSGQEQGH	359
DB	289	DVTHIAGSHDYTNMLASQWNSYDTEEPAPLGPDPFGKIQGLLT-----QTTIRAD	341
QY	360	YATRAHEAHDITDDPKVAPKGLTILKSG-SDDFNTNQPIRFTPVGM---GD---NNWRQ	412
DB	342	GSTRAHKATVSTGSHVHTPKLGSVQFTTDTNDFQAGQNTKFTPVGVIODGDHQNBPQQ	401
QY	413	WELPDYSGRLTLNNVLAPAVSPSPGRIILFFRSIVPSAGGSGYIDCLIPQEWVQHFI	472
DB	402	WSPNYSQRTGHNVLAPAVAPTFFGQLLFFRSTMPGCSGYNNMLDCLLQEWVLFHY	461
QY	473	QEMAPSQSAVALRVVNPDTGRNIFEAHLHREGFLTVANGNNPIVVPNGYFRFEAWGN	532
DB	462	QEMAPQSDVALLRFVNDTGRVLFECKLHKSQYITVAHTGPYDLVLPNGYFRFDSWN	521
QY	533	QFYTLAPMGSGQGRRA 549	
DB	522	QFYTLAPMGNGTGRRA 538	

RESULT 2

COAT_SOUV3	COAT_SOUV3	STANDARD;	PRT;	546 AA.
ID	COAT_SOUV3			
AC	Q04542;			
DT	01-OCT-1994 (Rel. 30, Created)			
DT	01-OCT-1994 (Rel. 30, Last sequence update)			
DT	01-OCT-1994 (Rel. 30, Last annotation update)			
DE	Coat protein (Capsid protein).			
OS	Southampton virus (serotype 3).			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;			
OC	Norovirus.			
NCBI_TaxID=37129;				
SEQUENCE FROM N.A.				
MEDLINE=93142023; PubMed=8380940;				
Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;				
"Sequence and genome organization of a human small round-structured				
(Norwalk-like) virus."				
Science 259:516-519 (1993).				
-1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.				
This SWISS-PROT entry is copyright. It is produced through a collaboration				
between the Swiss Institute of Bioinformatics and the EMBL outstation -				
the European Bioinformatics Institute. There are no restrictions on its				
use by non-profit institutions as long as its content is in no way				
modified and this statement is not removed. Usage by and for commercial				
entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
EMBL; L07418; AAA92984.1; -				
PIR; B37491; B37491.				
InterPro; IPR004005; Calici_coat.				
InterPro; IPR008975; Viral_cap_coat.				
Pfam; PF00915; Calici_coat; 1.				
Coat protein; Glycoprotein.				
CARBOHYD 73 73				
N-LINKED (GLCNAC. . .) (POTENTIAL).				
CARBOHYD 303 303				
N-LINKED (GLCNAC. . .) (POTENTIAL).				
CARBOHYD 340 340				
N-LINKED (GLCNAC. . .) (POTENTIAL).				
CARBOHYD 441 441				
N-LINKED (GLCNAC. . .) (POTENTIAL).				
SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;				

Query March 37.5%; Score 1089.5; DB 1; Length 546;

```

DR EMBL; M87481; AAA16217.1; -.
DR PIR; A48562; A48562.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 702 AA; 77850 MW; E6BSA58523DDEE3D7 CRC64;

Query Match 10.1%; Score 293; DB 1; Length 702;
Best Local Similarity 25.0%; Pred. No. 2.3e-14;
Matches 117; Conservative 61; Mismatches 150; Indels 140; Gaps 21;

QY 8 AAPSNDGAANLVPEANDEVMALPEVVGAS-----IAAPVVGQNNIIDPWIRENFVQAPQGE 63
DB 151 AESDGGGADIVTEEGTVVQOPVPAQSAITTLAASTG-----KTVDC 196
QY 64 FTV-----SPRNSPEMLLNLELGPENLYSLHSLRMVNGYAGGMQVQVVL 113
DB 197 WTTFFSYHTAVNWSITTEAQKILFSRALSPENLYSLHSLRSTWSGGIDVRFVSGSG 256
QY 114 FTAGKLIIPAAVPHPPFVVENISAAQITMCPHVIVDVROLEPVLPLPDIRNRFHYNQENT 173
DB 257 VFGGKLAALIVPP--GIEPVESPTMLQYPHVLFDAQTPEVIFTIPDIRKTLVH-SMDDT 313
QY 174 PMRLVAMLYTPI--RANGSEDVFTVSCVLRPADEPFEFTFLVPPTVESKTKPFLLPIL 231
DB 314 DITRLVIMYNELINPEYSEPKSCSIVETRPSSDFTFSLKLP--GSLLKHGSI 371
QY 232 TLGELS-----NSRFPAAIDMLYTDPNESIVQPO-----NGRCTLDGTLQGTQLVPTQIC 283
DB 372 LIPRNSRHMGNRWSTID-----GFVQPRVFSNRHDFDFTTIGWS--TPYYI- 420
QY 284 AFRGTLISQTARAADSTDSQRAARNHPLHVQKNLD-GTQYD-----P 325
DB 421 -----PIEVTLKLDKRGQYFKVTDTEKSLVPLGPDGWP 454
QY 326 TDDIPAVL-----GAIDFKGTGVGASQDVSGQEQGHVATRAHAHIOTDPKYAPKLG 381
DB 455 DTTIPTAMTASNGYDITVAEYRITN-----NGTHFKGYI-----MG 492
QY 382 TILIK-SGSDDF-NTNQPIR--FTPVGMGDNWNRQWELPDYSGRFLN 425
DB 493 NLTKVKYKSGDNLGETQTSRTLPASVG-----NYKQDNTIN 528

RESULT 4
POLN RHDV STANDARD; PRT; 2344 AA.
AC P27410;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Non-structural glycoprotein (Contains: RNA-directed RNA polymerase
DE (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like
DE protein); Coat protein)
OS Rabbit hemorrhagic disease virus (RHDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Lagovirus.
OC NCBI_TaxID=11976;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91361557; PubMed=1840711;
RA Meyers G., Wirblich C., Thiel H.-J.;
RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide
RT sequencing of a calicivirus genome.";
RL Virology 184:664-676 (1991).
CC -!- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE
CC FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +

```

```

[RNA] (N).
-!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC EMBL; M67473; AAA47285.1; -.
DR PIR; A41039; RWRWRH.
DR PDB; 1KHV; 16-JAN-02.
DR PDB; 1KHW; 16-JAN-02.
DR MEROPS; C24.001; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR000317; Peptidase_C24.
DR InterPro; IPR006055; RNA_helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR Pfam; PF03510; Peptidase_C24; 1.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00916; 2CENDOPTASE.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;
KW Thiol protease; Helicase; ATP-binding; Coat protein; 3D-structure.
FT CHAIN 2508 2707 HELICASE (P2C-LIKE).
FT CHAIN 1109 1251 PROTEASE P3C.
FT CHAIN 7149 71625 RNA-DIRECTED RNA POLYMERASE.
FT NP BIND 522 529 ATP (POTENTIAL).
FT ACT_SITE 1212 1212 PROTEASE (POTENTIAL).
FT ACT_SITE 1227 1227 PROTEASE (POTENTIAL).
SQ SEQUENCE 2344 AA; 257066 MW; 1454C248F81E9212 CRC64;

Query Match 10.0%; Score 290; DB 1; Length 2344;
Best Local Similarity 30.3%; Pred. No. 1.9e-13;
Matches 89; Conservative 49; Mismatches 114; Indels 42; Gaps 12;

QY 8 AAPSNDGAANLV-----VPEANDE-----VMALPEVV-----GASIAAPVVG-----QQNIID 49
DB 1772 AAPQGEAAGTATTASVPGTTDGMDFGVVAVTTSVITAENSSASIAATAGIGPPQQVDOOE 1831
QY 50 PWIRENFVQAPQGEFTVSPRNSPGEMLLNLELGPENLYSLHSLRMVNGYAGGMQVQVVL 109
DB 1832 TW-RTNFYY--NDVFTWSVADAPGSLITYVQHSPPNPPTAVLSQNYAGWAGMQRRPIV 1888
QY 110 AGNAFTAGKIIIPAAVPHPPFVVENISAAQITMCPHVIVDVROLEPVLPLPDIRNRFHYN 169
DB 1889 AGSGVFGGRLVRAVIPP--GIEIGPGLVVRQPHVVIDARSLEPVTITPDLRPNMYHT 1946
QY 170 QENTPRMLVAMLYTPIPLRANGSEDVFTVSCVLRPADEPFEFTFLVPP---TVESYTKP- 225
DB 1947 GDGGLVPTLVLSVYNNLINPFQGSTSAIQVTVETRPSEDFEFVIMRAPESKTVDSISPA 2006
QY 226 -FTLPILTLGELSNRFPAAIDMLYTDPNESIVQPOG---RCTLDTGLOGTT 275
DB 2007 LUTTPVLT-GVGNDRWNGQI-----VGLQPVPGFSTCNHWNHNGST 2049

RESULT 5
COAT_SMSV4
ID COAT_SMSV4 STANDARD; PRT; 703 AA.
AC P36285;
DT 01-JUN-1994 (Rel. 29, Created)

```

Wed Jun 2 09:13:24 2004

01-JUN-1994 (Rel. 29, Last sequence update)  
 15-DEC-1998 (Rel. 37, Last annotation update)  
 Coat protein (Capsid protein).  
 San Miguel sea lion virus (serotype 4) (SMV 4).  
 Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 Vesivirus.  
 NCBI\_TaxID=36407;  
 [1]  
 SEQUENCE FROM N.A.  
 MEDLINE=92410750; PubMed=1529644;  
 Neill J.D.;  
 "Nucleotide sequence of the capsid protein gene of two serotypes of  
 San Miguel sea lion virus: identification of conserved and non-  
 conserved amino acid sequences among calicivirus capsid proteins.";  
 Virus Res. 24:211-222(1992).  
 -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC  
 EMBL; M87482; AAA16220.1; -.  
 PIR; C48562; C48562.  
 InterPro; IPR004005; Calici\_coat.  
 InterPro; IPR008975; Viral\_cap\_coat.  
 Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 89 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 208 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 329 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 463 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 482 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 703 AA; 77721 MW; CSDAD823B261073 CRC64;  
 Query Match 9.9%; Score 287.5; DB 1; Length 703;  
 Best Local Similarity 24.5%; Pred. No. 5.9e-14;  
 Matches 104; Conservative 67; Mismatches 144; Indels 109; Gaps 18;  
 8 AAPSDNGAANLVPEANDEVMALEPVVG-----ASTAAPVVGQNIIDPWIRENFVQAPQGE 63  
 151 AESDGPGEAEIVTERQGTWQOQPAPAPATATLATATAGT-KSVEQEWTFYSHTSINW 209  
 64 FTVSPRNSGEMLLNLELGPENLPHSLSRMYNGYAGGMQVQVVLGNAGNAFTAGKIIFAA 123  
 210 STV---ESQKILYSQALNPSINPDLHIAKLYSTWGGIDVRFVTSVSGVFGGKLAALL 266  
 124 VPPHF-PVENISAAQITMCQPHVVDVRLQLEPVLLPLDIRNRPFFHYNQENTPRMLVAML 182  
 267 VPPGVEPIESVSMLOQ---PHVLFDAQVPEVIFTIPDIRKTLFH-SMDEDTTKLVINP 322  
 183 YTPLRANSEDEVFTVSCRVLTPAPDFETELVPP----- 217  
 323 Y-----ENGVENKTCISITVETPRASDFTFALLKPPGSLIKGSI1PSDLIPNSAHWMGNR 378  
 218 -----TVESKTKPFTLP--ILTLGEL-SNSRFPADIMLY 249  
 379 WNSTISGFSVQPRVFNQRHFDSTTTGWSSTPYVPIEIKIQGVGSKNKKFHVID--- 435  
 250 TD-----PNESIVQPNQGRCTLDGTLQ-GTTLVPTQICL-FRGTLIS 291  
 436 TDKALVPGIDGWPDTTIPDET---KATNGNFSGESYRAGSTTIKPNENSTHFKTYIC 492  
 292 QTARAADSDSPKARNRHPHVVQVKNLDGTQYDPTDDIPAVLGAIDFKGTGFGVASQRDV 351  
 493 GTLSTVETPEINDEQ-----QIKTEAEKSKQTMVVT-----ADFKDTI--VKPQHKI 537  
 352 SQQ 355  
 538 SPQK 541

RESULT 6  
 COAT\_FCVF4 STANDARD; PRT; 668 AA.  
 ID COAT\_FCVF4  
 AC P27405;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS Feline calicivirus (strain Japanese F4) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OK NCBI\_TaxID=11980;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91306470; PubMed=1853578;  
 RA Tohya Y., Taniiguchi Y., Takahashi E., Utagawa E., Takeda N.,  
 Miyamura K., Yamazaki S., Mikami T.;  
 "Sequence analysis of the 3'-end of feline calicivirus genome.";  
 Virology 183:810-814(1991).  
 CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 or send an email to license@isb-sib.ch).  
 CC  
 EMBL; D90357; BAA14371.1; -.  
 PIR; B40481; VCMWFC.  
 InterPro; IPR004005; Calici\_coat.  
 InterPro; IPR008975; Viral\_cap\_coat.  
 Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 399 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 668 AA; 73589 MW; 85BBD8CB85804E503 CRC64;  
 Query Match 9.7%; Score 283; DB 1; Length 668;  
 Best Local Similarity 24.2%; Pred. No. 1.2e-13;  
 Matches 131; Conservative 78; Mismatches 198; Indels 134; Gaps 27;  
 11 SNDGAANLVPEANDEV--MALEPVVGASIAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68  
 125 ADDGSIT-TPEQGTAVGGVIAEPPSAQMSTADMASGKSVDSM--EAFPSF-HTSVNWS 180  
 69 RNSPFGEMLLNLELGPENLPHSLSRMYNGYAGGMQVQVVLGNAGNAFTAGKIIFAAVPPHF 128  
 181 SETQGLKLFQSLGGLNPLNPHLEHLSKLYVWAGSIEVRSISGSGVFGGKLAIVVPPGV 240  
 129 -PVENISAAQITMCQPHVVDVRLQLEPVLLPLDIRNRPFFHYNQENTPRMLVAMLYTPL- 186  
 241 DPVQSTSMLOQ---PHVLFDAQVPEVIFTIPDIRKTLHV-MSDRTDTSLSVIMVNDLI 296  
 187 -----RANSGEDVFTVSCRVLTPAPDFETELVPPPTVESKTKPFTLPIITLIGELNSRF 241  
 297 NPYANDNSGSCIVTVE---TKPGDPFKFHLKPPG-----SVLTGHSI1PSDLI 342  
 242 PAALDM-----LYTDPNESIVQPNQGRCTLDGTLQGTTLVPTQICAPRGTLLISOTARA 296  
 343 PKSSSLWGNRYWTDITD-FVIRP-----FVFQANRH 373  
 297 AD---STDSPPQARNRHPHVVQVKNLDGTQYDPTDDIPAVLGAIDFKGTGFGVASQRDV 353  
 374 FDFNQETAGWSTPRFRPITITISEKNGSK-----LG-----IGVATDYIIP 415

```
QY 354 QOQGHYATRAHEAHDITDPKYPAPKL-----GTLIKSG-SDDPNTNQPIRFTPVGMGDDNN 409
DB 416 -----IPDGWPDFTT---IADKLIIPAGDYISITTEGNDIKTAQAYDTRAUVKNTTN 462
QY 410 WRQWELPDYSGRLTLNNNLAPAVSPFPGERILFFRSIVPSAGGYG-----SCYID---C 461
DB 463 FR-----GMYTCGSLQRAW-GDKKISNTAFITTAIRDNEIKPSNTIDMTKL 508
QY 462 LIPQEWQHFYQBAAPSQASVALRVYVNDPTGNIFEAKLHREGFLTV-----ANGGNP 516
DB 509 AVYQD--THVEQSVQTSDDTLALGYTG--IGEEAIGSNRDRVVRISVLPEAGARGGNHP 564
QY 517 I 517
DB 565 I 565

RESULT 7
COAT_FCVF9
ID COAT_FCVF9 STANDARD; PRT; 671 AA.
AC P27406;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 01-JUN-1994 (Rel. 29, Last annotation update)
DE Coat protein (Capsid protein).
GN CFP76.
OS Feline calicivirus (strain F9) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11981;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92410623; PubMed=1529544;
RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,
RA Turner P.C.;
RT "The complete nucleotide sequence of a feline calicivirus.";
RL Virology 190:443-448(1992).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339673; PubMed=1633955;
RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;
RT "Cloning and sequence determination of the feline calicivirus strain
RT F9.";
RL Biochem. Soc. Trans. 20:26S-26S(1992).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M86379; AAA79327.1; -.
DR EMBL; Z11536; CAA77636.1; -.
DR PIR; B43382; VCMWF9.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).
FT CARBOHYD 618 618 N-LINKED (GLCNAC. .) (POTENTIAL).
SQ SEQUENCE 671 AA; 73441 MW; 33BEE86D8370D5E5 CRC64;
Query Match 9.7%; Score 280.5; DB 1; Length 671;
Best Local Similarity 24.1%; Pred. No. 1.9e-13;
Matches 130; Conservative 74; Mismatches 207; Indels 129; Gaps 24;
QY 11 SNDGAANLVEANDEV--MALEFVVGASTAAPVVGQNIIDPWIRENFVQAPQGEFTVSP 68
```

```
DB 125 ADDGSIT-APEQGTWVGVIAPSPSAQMSTAADMATKSVDSSE--EAFPSF-HTSVNWST 180
QY 69 RNSPGMLNLNLGLPELNPYLSHLSRMVNGYAGGMQVQVVLVAGNAFTAGKIIPAAVPPHF 128
DB 181 SETQKILFKQSLGLPLNPYLEHLAKLYVANGSIEVRSISGSGVFGKLAIVVPPGV 240
QY 129 -PVENISAAQITCMCHVIVDVRLPEVLLPLDIRNRFPHYQENTPRMLVAMLYTPL- 186
DB 241 DPVQSTSMLOQ---PHVLFDAQVEVIFCLPDLRSTLVHL-MSDTRDTTSLVIMVNDLI 296
QY 187 -----RANGSEDVFTVSCRVLTFRPAPDPFEFTLVPTVESKYPFTLPILITGELSRSF 241
DB 297 NPYANDANSSGCIIVTVE---TKPGDPFKHLKPEG-----SMLTHGSIPSDLI 342
QY 242 PAA-----IDMLYTDPNESIVVQNGRCRTLDTGLQGTTLQVPTQICAFRGTLISQARAA 297
DB 343 PTKSSLWGNRYWSDITDFVIRP-----FVQANRHF 374
QY 298 D---STDSPQARNHPLHVQVKNLDGTQYDPTDDIPAVLGAIDPKGTVEGVASQRDVSGQ 354
DB 375 DFNQETAGWSTPRFRPISVTITEQNG-----AKLG-----IGVATDYIVPG- 415
QY 355 QEQGHYATRAHEAHDITDPKYPAPKLGTILIKSGS-DDPNTNQPIRFTPVGMGDDNNRWQW 413
DB 416 -----IPDGWPDFTTIPGELIPAGDYAITNGTNDITATTATYDGTADIIKNNTNFR-- 464
QY 414 ELPDYSGRLTLNNNLAPAVSPFPGERILFFRSIVPSA---GSGYSGYIDC-LIPOEWV- 468
DB 465 -----GMYTCGSLQRAW-GDKKISNTAFITTAITLDGNNKINPCNTIDDSKIV 512
QY 469 ----QHFOEAAAPSQASVALRVY-----VNPDTGRNIFEAKLHREGFLTVANGGNPFI 517
DB 513 VFQDNHVGKKAQTSDDTLALGYTGIGEOAIGSDRDRVVRISTLPETG---ARGGNHPI 568

RESULT 8
COAT_FCVF6
ID COAT_FCVF6 STANDARD; PRT; 668 AA.
AC P27404;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Coat protein (Capsid protein).
OS Feline calicivirus (strain CFI/68 FIV) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11979;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=91374597; PubMed=1716692;
RA Neill J.D., Reardon I.M., Heinrichson R.L.;
RT "Nucleotide sequence and expression of the capsid protein gene of
RT feline calicivirus.";
RL J. Virol. 65:5440-5447(1991).
RN [2]
RP SEQUENCE FROM N.A.
RA Neill J.D.;
RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
CC -!- PTM: The N-terminus is blocked.
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M32819; AAA42925.1; -.
DR EMBL; U13992; AAC13993.1; -.
DR PIR; A40507; VCMWFF.
```

empires requires a license (see [license@sib.ch](mailto:license@sib.ch)).

4



QY 543 -GGRR 547  
 DB 2200 MGRARR 2205

## RESULT 10

POLG\_FMDVO STANDARD; PRT; 2332 AA.  
 AC P03305;  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat  
 protein VP4; Coat protein VP2; Coat protein VP3; Coat protein VP1;  
 DE Core protein P12; Core protein P14; Core protein P14; Genome-linked  
 DE protein VP3; Protease (EC 3.4.22.-); RNA-directed RNA polymerase  
 DE (EC 2.7.7.48)].  
 OS Foot-and-mouth disease virus (strain O1) (Aphthovirus) (FMDV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Aphthovirus.  
 OX NCBI\_TaxID=73482;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=OLK;  
 RX MEDLINE=84297249; PubMed=6089122;  
 RA Fores S., Strebel K., Beck E., Schaller H.;  
 RT "Nucleotide sequence and genome organization of foot-and-mouth  
 RT disease virus";  
 RL Nucleic Acids Res. 12:6587-6601(1984).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O1BFS;  
 RX MEDLINE=83143292; PubMed=6298715;  
 RA Makoff A.J., Paynter C.A., Rowlands D.J., Boothroyd J.C.;  
 RT "Comparison of the amino acid sequence of the major immunogen from  
 RT three serotypes of foot and mouth disease virus";  
 RL Nucleic Acids Res. 10:8285-8295(1982).  
 RN [3]  
 RP X-RAY CRYSTALLOGRAPHY (2.9 ANGSTROMS).  
 RX MEDLINE=89143740; PubMed=2537470;  
 RA Acharya R., Fry E., Stuart D., Fox G., Rowlands D., Brown F.;  
 RT "The three-dimensional structure of foot-and-mouth disease virus at  
 RT 2.9-A resolution.";  
 RL Nature 337:709-716(1989).  
 CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC {RNA} (N).  
 CC -1- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC -1- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC -1- MISCELLANEOUS: THE STRAIN O1K SEQUENCE IS SHOWN.  
 CC -1- MISCELLANEOUS: THE COAT PROTEIN VP1 CONTAINS THE MAIN ANTIGENIC  
 CC DETERMINANTS OF THE VIRION; THEREFORE, CHANGES IN ITS SEQUENCE  
 CC MUST BE RESPONSIBLE FOR THE HIGH ANTIGENIC VARIABILITY OF THE  
 CC VIRUS.  
 CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X00871; CAA25416.1; -;  
 CC EMBL; J02185; AAA42635.1; -;  
 CC PDB; 1QWY; 18-SEP-01.  
 CC PDB; 1QOL; 10-NOV-00.  
 CC MEROPS; C03.008; -;  
 CC InterPro; IPR004004; Calici\_pol\_hel.  
 CC InterPro; IPR009003; Cys\_Ser\_trypsin.

DR InterPro; IPR004080; FMDVP1coat.  
 DR InterPro; IPR008739; Peptidase\_C28.  
 DR InterPro; IPR001676; Rhv.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_P5vir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF05408; Peptidase\_C28; 1.  
 DR Pfam; PF00073; rhv; 3.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR PRINTS; PRO0918; CALICVIRUSNS.  
 DR PRINTS; PRO1543; FMDVP1COAT.  
 KW Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;  
 KW Transferase; Hydrolase; Thiol protease; Nonstructural protein;  
 KW Myristate; 3D-structure; Lipoprotein.  
 FT CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.  
 FT CHAIN 202 286 COAT PROTEIN VP4.  
 FT CHAIN 287 504 COAT PROTEIN VP2.  
 FT CHAIN 505 724 COAT PROTEIN VP3.  
 FT CHAIN 725 937 COAT PROTEIN VP1.  
 FT CHAIN 938 1107 CORE PROTEIN P12.  
 FT CHAIN 1108 1425 CORE PROTEIN P13.  
 FT CHAIN 1426 1578 CORE PROTEIN P14.  
 FT CHAIN 1579 1649 CORE PROTEIN P14.  
 FT CHAIN 1650 1862 GENOME-LINKED PROTEIN VP3.  
 FT CHAIN 1863 2332 RNA-DIRECTED RNA POLYMERASE.  
 FT CHAIN 202 202 N-myristoyl glycine (in host).  
 FT LIPID 511 511 INTERCHAIN (IN VP3 DIMER).  
 FT DISULFID 406 858 IN VP2-VP1 DIMER.  
 FT DISULFID 780 780 I -> V (IN STRAIN O1BFS).  
 FT VARIANT 808 808 G -> R (IN STRAIN O1BFS).  
 FT VARIANT 861 861 N -> S (IN STRAIN O1BFS).  
 SQ SEQUENCE 2332 AA; 258924 MW; 4A83176F43447D68 CRC64;

Query Match 4.5%; Score 129.5; DB 1; Length 2332;

Best Local Similarity 21.2%; Pred. No. 0.24;

Matches 75; Conservative 35; Mismatches 131; Indels 113; Gaps 15;

QY 79 LEIGPELNYLSHLSRMVYAGGQVOVLVLAAGNAFTAGKIIFAAVPPHPPVENISAQI 138  
 DB 367 LELPTDHKVGSLTDSYAMRNGWDVEVTANGNQNGGCLLVAMPYELYSIOKRELYQL 426  
 QY 139 TMCPHVIVDRQLEPVLPLPDIR-NRFFHYQENCTPRMRLVAMLYTPLRANSGEDVFTV 197  
 DB 427 TLFPHQFINPRNTMTAHITVPFVGNRYDQYKHKP--WTLVVMVVAPLTVN----- 476  
 QY 198 SCRVLTRPADPEFTFLVPPTVESKTKPTLTLTLGELSISR--PPAID-----MLYT 250  
 DB 477 -----TEGAQIKVYVANIAPT-----NHHVAGEPPSKEGIFPVACSDGYGLVTT 521  
 QY 251 DPNES-----IVVQPN---GRCITLDTGTTQLVPTQICAFRCITLISQTAARAADSTD 301  
 DB 522 DKPTADPVYGVKFNPPRNQLPGRFT---NLLDVAEACPT-FLRFEQGVVYVTK-----TD 573  
 QY 302 SPQARNHPLHVQVKNLDG-----TQYD-----PTD----- 327  
 DB 574 SRVLAAQPDMSLAQKMSNTFLAGLAQYVTVQSGTINLHFMFTGPTDAKARYMVAVPPG 633  
 QY 328 -----DIPAVLGAIDFKGTGVGVASQDVSG 353  
 DB 634 MBPPKTPAAAHCIHAEMDTGLNSKPTFSIP-YLSAADYVAYTAGSVAAETTNVQG 686

## RESULT 11

POLG\_HE701  
 ID POLG HE701 STANDARD; PRT; 2194 AA.  
 AC P32537;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2

DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein  
DE P2A; Core protein P2B; Core protein P2C; Core protein P2A; Genome-  
DE linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
DE (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)]  
OS Human enterovirus 70 (strain J670/71) (EV 70).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Enterovirus.  
OX NCBI\_TaxID=31915;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=91037960; PubMed=2172447;  
RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D.,  
RA Minor P.D., Almond J.W.;  
RT "The complete nucleotide sequence of enterovirus type 70:  
RT relationships with other members of the picornaviridae.";  
RL J. Gen. Virol. 71:2291-2299(1990).  
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.  
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the  
CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
CC substituted for Gln, and Ser or Thr for Gly.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC (RNA) (N).  
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
CC each of which is composed of one copy each of proteins VP1, VP2,  
CC VP3, and VP4.  
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC -----  
CC EMBL: D00820; BAAL8991.1; -;  
CC FIR: A36253; GNNYE7.  
CC HSP: P03300; IPOV.  
CC MEROPS: C03\_UPE; -;  
CC MEROPS: C03\_UPE; -;  
CC InterPro: IPR003593; AAA\_ATPase.  
CC InterPro: IPR004004; Calici\_pol\_hel.  
CC InterPro: IPR009003; Cys\_Ser\_trypsin.  
CC InterPro: IPR000199; Pept\_3C\_picorn.  
CC InterPro: IPR000081; Peptidase\_C3.  
CC InterPro: IPR003138; Pico\_P1A.  
CC InterPro: IPR002527; Pico\_P2B.  
CC InterPro: IPR001676; Rhv.  
CC InterPro: IPR006055; RNA\_helicase.  
CC InterPro: IPR007095; RNA\_pol\_DS\_PS.  
CC InterPro: IPR001205; RNA\_pol\_P3D.  
CC InterPro: IPR007094; RNA\_pol\_PSVir.  
CC InterPro: IPR008975; Viral\_cap\_coat.  
CC Pfam: PF00548; Cys-protease\_3C; 1.  
CC Pfam: PF02226; Pico\_P1A; 1.  
CC Pfam: PF00947; Pico\_P2A; 1.  
CC Pfam: PF01552; Pico\_P2B; 1.  
CC Pfam: PF00073; rhv; 3.  
CC Pfam: PF00680; RNA\_dep\_RNA\_pol; 1.  
CC Pfam: PF00910; RNA\_helicase; 1.  
CC PRINTS: PR00918; CALICVIRUSNS.  
CC ProDom: PD001125; Cys\_protease\_3C; 1.  
CC ProDom: PD001306; Pico\_P2A; 1.  
CC ProDom: PD001274; Pico\_P2B; 1.  
CC SMART: SM00382; AAA; 1.  
KW Polyprotein; Core protein; Transferase;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
KW Lipoprotein.  
FT CHAIN 2 69 COAT PROTEIN VP4.  
FT CHAIN 70 319 COAT PROTEIN VP2.  
FT CHAIN 320 561 COAT PROTEIN VP3.

FT CHAIN 871 COAT PROTEIN VP1.  
FT CHAIN 872 CORE PROTEIN P2A.  
FT CHAIN 1014 CORE PROTEIN P2B.  
FT CHAIN 1113 CORE PROTEIN P2C.  
FT CHAIN 1114 CORE PROTEIN P2A.  
FT CHAIN 1443 CORE PROTEIN P2C.  
FT CHAIN 1444 CORE PROTEIN P2A.  
FT CHAIN 1532 GENOME-LINKED PROTEIN VPG.  
FT CHAIN 1533 GENOME-LINKED PROTEIN VPG.  
FT CHAIN 1534 PICORNAIN 3C.  
FT CHAIN 1535 RNA-DIRECTED RNA POLYMERASE P3D.  
FT CHAIN 1737 N-myrhistoyl glycine (in host) (By  
FT CHAIN 1738 similarity)  
FT CHAIN 1739 PROTEASE (POTENTIAL).  
FT ACT\_SITE 1701 PROTEASE (POTENTIAL).  
FT ACT\_SITE 1715 PROTEASE (POTENTIAL).  
FT ACT\_SITE 1715 PROTEASE (POTENTIAL).  
SQ SEQUENCE 2194 AA; 244590 MW; 15DBAB96EE06673C CRC64;  
Query Match 4.3%; Score 125.5; DB 1; Length 2194;  
Best Local Similarity 23.3%; Pred. No. 0.44;  
Matches 55; Conservative 35; Mismatches 85; Indels 61; Gaps 10;  
OY 66 VSPRNSPGEMLLNLEL-----GPELNPYLSHLSRMVNGYAGQVQVVLGNAFTAGKII 120  
DB 390 ISAGSDMDQLLEFIPLDIOLEGLPLRTLLGNISRYVTHWSGLEWTFMFCGSMFTTKLI 449  
OY 121 FAAVPP--HPFVENISAAQITMCPHVIVDVRLQEPVLLPLPDIRNRRFFHYNOENTPRML 178  
DB 450 ICYTPPGSGSPTDRMQAMLAT---HVWDFGLSSITIIIPWISGS--HYRMENTDKAI 504  
OY 179 VA-----MLYPLRANS--EDVFTVSCR-----LTPADDFEFTELVP-- 217  
DB 505 NANVGYTCFQTNLAVPVGAAQCIVGVAAKDFNLMLRSDPDIGOSAILPQMAT 564  
OY 218 -----TVESKTKP-----TLPILTIGELSNRFPAAIDMLYTDPNESI 256  
DB 565 TQIGEIVKTVANTVESEIKAEGLVIPSINAVEIGATSN-----TEPBEAI 609  
RESULT 12  
POLG\_FMDV1 STANDARD; PRT: 2333 AA.  
ID POLG\_FMDV1 STANDARD; PRT: 2333 AA.  
AC P03306; Q84750; Q84751; Q84752; Q84753; Q84754; Q84760;  
AC Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;  
AC Q84769; Q89824;  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat  
DE proteins VP1 TO VP4; Core protein P52; Genome-linked proteins VPG1 TO  
DE VP3; Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed  
DE RNA polymerase P56A (EC 2.7.7.48)].  
DE Foot-and-mouth disease virus (strain A10-61) (Aphthovirus A) (FMDV).  
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Aphthovirus.  
OC NCBI\_TaxID=12112;  
OX [1]  
RX SEQUENCE FROM N.A.  
RP MEDLINE=84169547; PubMed=6324120;  
RX Carroll A.R., Rowlands D.J., Clarke B.E.;  
RA "The complete nucleotide sequence of the RNA coding for the primary  
RA translation product of foot and mouth disease virus.";  
RT Nucleic Acids Res. 12:2461-2472(1984).  
RL [2]  
RP SEQUENCE OF 115-1048 FROM N.A.  
RX MEDLINE=82211814; PubMed=6282711;  
RA Boothroyd J.C., Harris T.J.R., Rowlands D.J., Lowe P.A.;  
RA "The nucleotide sequence of cDNA coding for the structural proteins  
RT of foot-and-mouth disease virus.";  
RL Gene 17:153-161(1992).  
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-|-Gly bond in the  
CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
CC substituted for Gln, and Ser or Thr for Gly.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
CC (RNA) (N).  
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
CC each of which is composed of one copy each of proteins VP1, VP2,  
CC VP3, and VP4.



RX MEDLINE=20556633; PubMed=1102480;  
 RA Petrecca K., Miller D.M., Shrier A.;  
 RT "Localization and enhanced current density of the Kv4.2 potassium  
 RL channel by interaction with the actin-binding protein filamin.";  
 RL J. Neurosci. 20:8736-8744(2000).  
 RL [8]  
 RN INTERACTION WITH INPPL1.  
 RP MEDLINE=21601813; PubMed=11739414;  
 RX Dyson J.M., O'Malley C.J., Becanovic J., Munday A.D., Berndt M.C.,  
 RA Coghill I.D., Nandurkar H.H., Ooms L.M., Mitchell C.A.;  
 RT "The SH2-containing inositol polyphosphate 5-phosphatase, SHIP-2,  
 RL binds filamin and regulates submembraneous actin.";  
 RL J. Cell Biol. 155:1065-1079(2001).  
 RL [9]  
 RN REVIEW.  
 RP MEDLINE=21234905; PubMed=11336782;  
 RX van der Flier A., Sonnenberg A.;  
 RA "Structural and functional aspects of filamins.";  
 RL Biochim. Biophys. Acta 1538:99-117(2001).  
 RL [10]  
 RN REVIEW.  
 RP MEDLINE=21146932; PubMed=11252955;  
 RX Stossek T.P., Condeelis J., Cooley L., Hartwig J.H., Noegel A.,  
 RA Schleicher M., Shapiro S.S.;  
 RT "Filamins as integrators of cell mechanics and signalling.";  
 RL Nat. Rev. Mol. Cell Biol. 2:138-145(2001).  
 RL [11]  
 RN VARIANTS PH PHE-656 AND THR-1764.  
 RP MEDLINE=21423782; PubMed=11532987;  
 RX Duncan J.S., Dixon P.H., Fox J.W., Hong S.E., Kinton L., Sisodiya S.M.,  
 RA Sheen V.L., Dubeau F., Scheffer I.E., Schachter S.C., Wilner A.,  
 RA Henchy R., Crino P., Kamuro K., DiMario F., Berg M., Kuzniick R.,  
 RA Cole A.J., Bromfield E., Biber M., Schomer D., Wheless J., Silver K.,  
 RA Mochida G.H., Berkovic S.F., Andermann F., Andermann E., Dobyns W.B.,  
 RA Wood N.W., Walsh C.A.;  
 RT "Mutations in the X-linked filamin 1 gene cause periventricular  
 RL nodular heterotopia in males as well as in females.";  
 RL Hum. Mol. Genet. 10:1775-1783(2001).  
 RL [12]  
 RN VARIANT PH VAL-82.  
 RP MEDLINE=21912461; PubMed=11914408;  
 RX Moro F., Carozzo R., Veggliotti P., Tortorella G., Toniolo D.,  
 RA Volzone A., Guerrini R.;  
 RT "Familial periventricular heterotopia: missense and distal truncating  
 RL mutations of the FLN1 gene.";  
 RL Neurology 58:916-921(2002).  
 RL [13]  
 RN VARIANTS OPD1 PHE-172; TRP-196 AND LEU-207, VARIANTS OPD2 PRO-170;  
 RP GLY-196; SER-200; LYS-254; PRO-273; LYS-555 AND PHE-1645, VARIANTS FMD  
 RA ALA-1159; LEU-1186 AND ILE-1620 DEL, AND VARIANTS MNS GLU-1184;  
 RP THR-1188 AND LEU-1199.  
 RX MEDLINE=22552708; PubMed=12612583;  
 RA Robertson S.P., Twigg S.R.F., Sutherland-Smith A.J., Biancalana V.,  
 RA Gorlin R.J., Horn D., Kenwick S.J., Kim C.A., Morava E.,  
 RA Newbury-Ecob R., Oerstavik K.H., Quarrell O.W.J., Schwartz C.E.,  
 RA Shears D.J., Suri M., Kendrick-Jones J., Wilkie A.O.M.;  
 RT "Localized mutations in the gene encoding the cytoskeletal protein  
 RL filamin A cause diverse malformations in humans.";  
 RL Nat. Genet. 33:487-491(2003).  
 CC -1- FUNCTION: Promotes orthogonal branching of actin filaments and  
 CC links actin filaments to membrane glycoproteins. Anchors various  
 CC transmembrane proteins to the actin cytoskeleton and serves as a  
 CC scaffold for a wide range of cytoplasmic signaling proteins.  
 CC Interaction with FLNA may allow neuroblast migration from the  
 CC ventricular zone into the cortical plate. Tethers cell surface-  
 CC localized furin, modulates its rate of internalization and directs  
 CC its intracellular trafficking (By similarity).  
 CC -1- SUBUNIT: Homodimer. Interacts with FLNB, FURIN, HSPB7, INPPL1,  
 CC PSEN1 and PSEN2. Interacts also with various other binding  
 CC partners in addition to filamentous actin. Interacts  
 CC with KCND2.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic; peripheral.  
 CC -1- TISSUE SPECIFICITY: Ubiquitous.

CC -1- DOMAIN: Comprised of a NH2-terminal actin-binding domain, 24  
 CC internally homologous repeats and two hinge regions. Repeat 24 and  
 CC the second hinge domain are important for dimer formation.  
 CC -1- PTM: Phosphorylation extent changes in response to cell  
 CC activation.  
 CC -1- PTM: The N-terminus is blocked.  
 CC -1- DISEASE: Defects in FLNA are the cause of periventricular  
 CC heterotopia (PH) [MIM:300049]; also called nodular heterotopia,  
 CC bilateral periventricular (NBHP or BNPH). PH is an X-linked  
 CC developmental dominant disorder in which many neurons fail to  
 CC migrate into the cerebral cortex. They remain as nodules lining  
 CC the ventricular surface. In heterozygous females these neurons  
 CC presumably represent those cells that, after X-chromosome  
 CC inactivation, contain the active X chromosome with the filamin  
 CC mutation. Most hemizygous affected males die early during  
 CC embryogenesis, whereas heterozygous females have normal  
 CC intelligence but suffer from seizures and various manifestations  
 CC outside the central nervous system, especially related to the  
 CC vascular system. This implies that essential embryonic cell  
 CC migration can only occur in FLNA-expressing cells.  
 CC -1- DISEASE: Defects in FLNA are the cause of otopalatodigital  
 CC syndrome type I (OPD1) [MIM:311300].  
 CC -1- DISEASE: Defects in FLNA are the cause of otopalatodigital  
 CC syndrome type II (OPD2) [MIM:304120]; also known as  
 CC cranioodigital syndrome. OPD2 is a congenital bone disorder that  
 CC is characterized by abnormally modeled, bowed bones, small or  
 CC absent first digits and, more variably, cleft palate, posterior  
 CC fossa brain anomalies, omphalocele and cardiac defects.  
 CC -1- DISEASE: Defects in FLNA are the cause of frontometaphyseal  
 CC dysplasia (FMD) [MIM:305620]. FMD is a congenital bone disease  
 CC characterized by supraorbital hyperostosis, deafness and digital  
 CC anomalies.  
 CC -1- DISEASE: Defects in FLNA are the cause of Melnick-Needles syndrome  
 CC (MNS) [MIM:309350]. MNS is a severe congenital bone disorder  
 CC characterized by typical facies (exophthalmos, full cheeks,  
 CC micrognathia and malalignment of teeth), flaring of the metaphyses  
 CC of long bones, S-like curvature of bones of legs, irregular  
 CC contractions in the ribs, and sclerosis of base of skull.  
 CC -1- SIMILARITY: Belongs to the filamin family (CH) domains.  
 CC -1- SIMILARITY: Contains 2 calponin-homology  
 CC -1- SIMILARITY: Contains 24 filamin repeats.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 CC EMBL; X53416; CAA37495.1; -;  
 CC EMBL; L44140; AAA92644.1; -;  
 CC EMBL; X70082; CAA49687.1; -;  
 CC EMBL; X70085; CAA49690.1; -;  
 CC PIR; A37098; A37098.  
 CC HSPB; P13466; IKSX.  
 CC Genew; HGNC:3754; FLNA.  
 CC MIM; 300017; -;  
 CC MIM; 300049; -;  
 CC MIM; 304120; -;  
 CC MIM; 305620; -;  
 CC MIM; 309350; -;  
 CC MIM; 311300; -;  
 CC -----

Query Match 4.1%; Score 119; DB 1; Length 2647;  
 Best Local Similarity 21.0%; Pred. No. 1.8;  
 Matches 131; Conservative 64; Mismatches 219; Indels 209; Gaps 31;  
 48 IDPWREN-----FVQAPQ-GEFTVSPR-----NSPCEMLNLGPELNPYL--S 90  
 1696 VDVDVVENEDGTDFIDFYTAQPGKVICVRFGEHVPSPFQVFTALAGDQFSVQPPLRSQ 1755  
 91 HLSRMVNGYAGGMQVQVVLGNAGTAKIIFAAVPPHPPVENISAAQITMCPHVIVDVQR 150

```

Db 1756 QLAPOQTYAGGQOTWA-----PERPLVGNG-----LDVTS 1787
QY 151 LEPVLLPLDIRRPFPHYQNTPRML-----VAMLYTPLRANSGB- 192
Db 1788 LRPFDLVIP-----FTIKGEITGEVRMPSGKVAQPTITDNKDGTVTVYARSEAGLHEM 1842
QY 193 -----DVTVCRLVTRPAPDFEFTFLVPPTVESKTKPTFLPILTLGEL 236
Db 1843 DIRYDNMHIPGSPLOQYVDYVNGGHVTA YCPG-----LTHGVNKPATFTVNTKDAGSG 1896
QY 237 SNS---RFPAAIDMLYTDNE-----SIVVQ-----PQNGRCTLG 269
Db 1897 GLSLAIEGSPSKAEISCTNQDGTCSVSLPVLPGDYSLVLYKNEQHVPGSPPTARVTGDD 1956
QY 270 TLQ-----GTTQLVPTQICAFRGTLISQARAADSTDSF---QRARN----- 308
Db 1957 SMRSHLKVGSAAIDIPINSETDLSLTATVVPSPGSEPCLLKRLRNHGVHSFVPKET 2016
QY 309 --HPLHVQVKNLDGTQYDPTDIPAVLGAIDFKGTGVFGVASQDVSQO-QEQQHYATRAH 365
Db 2017 GEHLVHVK-KN---GQHVASSPIPVVISQSE-----IGDASKRVVSGQGLHEGHTPEPA- 2066
QY 366 EAHIDTDPKYAPKLGTLIKSGSD--DENTNQPIRFTVPVGMDDNNRWQWELPDYSGRLT 423
Db 2067 EFTIDTRDAGY---GLSLIEGSPSKVDINTED-----LEDGTCRVTYCTPTEFGNYI 2115
QY 424 LNMNLAPAVSPSPF-----GE-----RILFFRSIVPSAGGYG-IDLCLIPQEWQHF- 471
Db 2116 INIKFADQHVGPSFSPKVTGGRVKESITRRRAPSVANVSGHCDLSLKIPEISIQDWT 2175
QY 472 YQEAAPS-----QSAVALRVYNPDTGRNIFEAKLHREGFLTVANGCNPVIVP 520
Db 2176 AQVTSBPGKTHEAEIVEGENHTYCFIRFVPAENGTHTVSVKYKGQHV----- 2221
QY 521 PNGYFRFEAGNQFYTLAPMGSG 543
Db 2222 PGSPFQF-----TVGPLGEG 2236

```

## RESULT 14

```

POLG_CX16T STANDARD; PRT; 2193 AA.
ID Q9QF31;
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2
DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein
DE P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-
DE linked protein VPG (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)
DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Coxsackievirus A16 (strain Tainan/5079/98).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=231417;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21427116; PubMed=11536241;
RA Yan J.-J., Su I.-J., Chen P.-F., Liu C.-C., Yu C.-K., Wang J.-R.;
RT "Complete genome analysis of enterovirus 71 isolated from an outbreak
RT in Taiwan and rapid identification of enterovirus 71 and
RT coxsackievirus A16 by RT-PCR.";
RL J. Med. Virol. 65:331-339 (2001).
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular
CC membranes and is associated with viral RNA synthesis.
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
CC Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
CC poliovirus polyprotein. In other picornavirus reactions Glu may be
CC substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).

```

```

CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
CC each of which is composed of one copy each of proteins VP1, VP2,
CC VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF177911; AAC55085.1; -.
CC HSP; P03300; IPOV.
CC InterPro: IPR003593; AAA_ATPase.
CC InterPro: IPR004004; Calci_pol_hel.
CC InterPro: IPR009003; Cys_Ser_trypsin.
CC InterPro: IPR001099; Pept_3C_picorn.
CC InterPro: IPR000081; Peptidase_C3.
CC InterPro: IPR003138; Pico_P1A.
CC InterPro: IPR002527; Pico_P2B.
CC InterPro: IPR001676; Rhv.
CC InterPro: IPR006005; RNA_helicase.
CC InterPro: IPR007095; RNA_pol_DS_PS.
CC InterPro: IPR001205; RNA_pol_P3D.
CC InterPro: IPR007094; RNA_pol_PSVir.
CC InterPro: IPR008975; Viral_cap_coat.
CC Pfam; PF00548; Cys-protease_3C; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF00073; rhv_3.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICVIRUSNS.
CC ProDom; PD001125; Cys_protease_3C; 1.
CC ProDom; PD001306; Pico_P2A; 1.
CC ProDom; PD001274; Pico_P2B; 1.
CC SMART; SM00382; AAA; 1.
CC Polyprotein; Coat protein; Core protein; Transferrase;
CC RNA-directed RNA polymerase; Hydrolyase; Thiol protease; Myristate;
CC Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 323 COAT PROTEIN VP2.
FT CHAIN 324 565 COAT PROTEIN VP3.
FT CHAIN 566 862 COAT PROTEIN VP1.
FT CHAIN 863 1012 CORE PROTEIN P2A.
FT CHAIN 1013 1111 CORE PROTEIN P2B.
FT CHAIN 1112 1440 CORE PROTEIN P2C.
FT CHAIN 1441 1526 CORE PROTEIN P3A.
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1549 1731 PICORNAIN 3C.
FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.
FT LIPID 2 2 N-myristoyl glycine (in host) (By similarity).
FT ACT_SITE 1695 1695 PROTEASE (POTENTIAL).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
SQ SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;
Query Match 4.1%; Score 118.5; DB 1; Length 2193;
Best Local Similarity 32.5%; Pred. No. 1.5;
Matches 37; Conservative 15; Mismatches 49; Indels 13; Gaps 4;
QY 57 VQAPGGEFTVSRNSPFGMLNLELGPENLPYLSHLSRMVNYAGMQVQVVLGNAFTA 116
Db 398 VQSKTGELCAAPRADPGRD-----GPMQSTILGLCRYTQWSGSLEVTMFAGSFMAT 451
QY 117 GXIIIAAVPP--HFPVENISAQITMCHVIVDVQLEPVLPLPDIRNRFHY 168
Db 452 GKMLIAYTPPGGNVPADRITAMLGT---HVIDWDFGLQSSVTLVWPWISNT--HY 500

```

FT CHAIN 1427 1579 CORE PROTEIN P19.  
 FT CHAIN 1580 1602 GENOME-LINKED PROTEIN VP1.  
 FT CHAIN 1603 1626 GENOME-LINKED PROTEIN VP2.  
 FT CHAIN 1627 1650 GENOME-LINKED PROTEIN VP3.  
 FT CHAIN 1651 1863 PROTEASE.  
 FT CHAIN 1864 2336 RNA-DIRECTED RNA POLYMERASE.  
 FT CHAIN 202 N-methylglycine (in host) (By similarity).  
 FT LIPID 202  
 SQ SEQUENCE 2336 AA; 259982 MW; 15AC2AB022B5B954 CRC64;  
 Query Match 4.1%; Score 118; DB 1; Length 2336;  
 Best Local Similarity 19.0%; Pred. No. 1.8; Indels 82; Gaps 8;  
 Matches 67; Conservative 46; Mismatches 158;  
 Qy 64 FTVSPRNSPCGEMLLNLELGPENLPYLHLRLVYAGVQVQVVLGNAFTAGKIIFAA 123  
 Db 353 FDMTPDKAFGH-LEKLELPTDHHKGVYGHVDSFAYMENGWDEVSAGVQNGGCLLVAM 411  
 Qy 124 VPHFPVENISAAQITMCPHVIIVDRQLBEPVLLPLDIRRPFHYNQENTPSMRLVAMLY 183  
 Db 412 VPEWKELTPREKYQLTLFPHQFI-----SPRTNMTAHIV 445  
 Qy 184 TPLRANGSEDFVTVSCRVLTREPAPDEFPLVPTVESKTKPTLPLTLGELNSRFP 243  
 Db 446 VPYLGAVRYDQY-----KKHKPWLVMVVSPLTTVTSA 480  
 Qy 244 AIDMLYTDPNESIVVQPNQGRCTLDGTLOGTTLQVPTQICAFRGTLISQTAARADSDSP 303  
 Db 481 QQIKVYAN-----IAPH--VHAGELPSKEGIVPVACSDGYGLVTTDPKTAD----- 527  
 Qy 304 QRARNHPLHVQVKNLDGTQY-----DPTDDIPAVLGAIDPKGTGVGVASQRDVSGQ 355  
 Db 528 -----PVYGMVYNPPRTNYPGRFTNLLDVAEACPTFLCFDDGKPYVVTDTDEQLLAKF 581  
 Qy 356 EQGHVATTAHEAHDITTDPKYAPKLGTI---LIKSGSDDFNNTNPIRFTPVGM 405  
 Db 582 DLJSLAAKMSNTYLSGIAQYQAQYSGTINLHFMFTGSTDSKARYWVAVYPPGV 634

Search completed: June 1, 2004, 13:47:25  
 Job time : 8.64429 secs

RESULT 15  
 POLG\_FMDVZ STANDARD; PRT; 2336 AA.  
 AC P49303;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Genome polyprotein [Contains: Nonstructural protein P20A; Coat  
 DE proteins VP1 TO VP4; Core proteins X, P14, P41, P19; Genome-linked  
 DE proteins VP1 TO VP3; Picornain 3C (EC 3.4.22.28) (Protease 3C)  
 DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48)].  
 OS Foot-and-mouth disease virus (strain A22/550 Azerbaijan 65)  
 OS (Aphthovirus A) (FMDV).  
 OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
 OC Aphthovirus.  
 OC NCBI\_TaxID=73481;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Sosnovtsev S.V., Onischenko A.M., Petrov N.A., Kalashnikova T.I.,  
 RA Mamaeva N.V., Drygin V.Y., Perevozchikova N.A., Vasilenko S.K.;  
 RA Submitted (AUG-1993) to the EMBL/GenBank/DBJ databases.  
 RL CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the  
 CC poliovirus polyprotein. In other picornavirus reactions Glu may be  
 CC substituted for Gln, and Ser or Thr for Gly.  
 CC CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
 CC [RNA] (N).  
 CC SUBUNIT: The virus capsid is composed of 60 icosahedral units,  
 CC each of which is composed of one copy each of proteins VP1, VP2,  
 CC VP3, and VP4.  
 CC PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
 CC SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
 CC  
 CC The SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X74812; CAA52812.1; --  
 DR PIR; S37077; S37077.  
 DR HSP; O88571; 1TME.  
 DR MEROPS; C03.008; --  
 DR MEROPS; C28.001; --  
 DR InterPro; IPR004004; Calici\_pol\_hel.  
 DR InterPro; IPR009003; Cys\_ser\_trypsin.  
 DR InterPro; IPR004080; FMDVPicoat.  
 DR InterPro; IPR008739; Peptidase\_C28.  
 DR InterPro; IPR001676; Rhv.  
 DR InterPro; IPR000605; RNA\_helicase.  
 DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
 DR InterPro; IPR001205; RNA\_pol\_P3D.  
 DR InterPro; IPR007094; RNA\_pol\_P5vir.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF05408; Peptidase\_C28; 1.  
 DR Pfam; PF00073; rhv; 3.  
 DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
 DR Pfam; PF00910; RNA\_helicase; 1.  
 DR Pfam; PF00918; CALICVIRUSNS.  
 DR PRINTS; PR01542; FMDVPICOAT.  
 DR Polyprotein; Coat protein; Core protein; RNA-directed RNA polymerase;  
 DR Transferase; Hydrolase; Thiol protease; Nonstructural protein;  
 DR Myristate; Lipoprotein.  
 KW CHAIN 1 201 NONSTRUCTURAL PROTEIN P20A.  
 KW CHAIN 202 286 COAT PROTEIN VP4.  
 FT CHAIN 287 504 COAT PROTEIN VP2.  
 FT CHAIN 505 724 COAT PROTEIN VP3.  
 FT CHAIN 725 938 COAT PROTEIN VP1.  
 FT CHAIN 939 954 CORE PROTEIN X.  
 FT CHAIN 955 1108 CORE PROTEIN P14.  
 FT CHAIN 1109 1426 CORE PROTEIN P41.

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 32.0218 Seconds  
(without alignments)  
5419.275 Million cell updates/sec

Title: US-09-926-799-10  
Perfect score: 2906  
Sequence: 1 MKMASNDAAPSNDGAANLVP.....GNQFYTLAPMGSGQGRRAQ 550

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2906	100.0	550	12 Q8JW42	Q8JW42 norwalk-lik
2	2861	98.5	550	12 Q318B6	Q318B6 norwalk-lik
3	2855	98.2	550	12 Q318B3	Q318B3 norwalk-lik
4	2847	98.0	550	12 Q91V37	Q91V37 human calic
5	2828	97.3	550	12 Q8V778	Q8V778 norwalk vir
6	2819	97.0	550	12 Q8V781	Q8V781 norwalk vir
7	2637.5	90.8	547	12 Q318A7	Q318A7 norwalk-lik
8	2630.5	90.5	547	12 Q8V775	Q8V775 norwalk vir
9	2623.5	90.3	547	12 Q8V773	Q8V773 norwalk vir
10	2129	73.3	548	12 Q66296	Q66296 calicivir
11	2129	73.3	548	12 Q8V768	Q8V768 norwalk vir
12	2126	73.2	548	12 Q8V0P5	Q8V0P5 human calic
13	2121	73.0	548	12 Q8V0P4	Q8V0P4 human calic
14	2117	72.8	548	12 Q317Z2	Q317Z2 norwalk-lik
15	2117	72.8	548	12 Q317Z5	Q317Z5 norwalk-lik
16	2116	72.8	548	12 Q318A1	Q318A1 norwalk-lik

ALIGNMENTS

RESULT 1

Q8JW42	PRELIMINARY;	PRT;	550 AA.
ID	Q8JW42		
AC	Q8JW42;		
DC	01-OCT-2002 (Tremblrel. 22, Created)		
DT	01-OCT-2002 (Tremblrel. 22, Last sequence update)		
DT	01-OCT-2003 (Tremblrel. 25, Last annotation update)		
DE	Capsid protein.		
GN	CAPSID.		
OS	Norwalk-like virus.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_TaxID=95340;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Ueno7k;		
RA	Katayama K., Takeda N., Natori K.;		
RL	Submitted (JAN-2002) to the EMBL/GenBank/DDBJ databases.		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Ueno7k;		
RA	Natori K., Takeda N.;		
RT	"Genetic and antigenic relationship among Norwalk-like viruses.";		
RL	Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.		
DR	EMBL, AB078337; BAC05518.1;		
DR	InterPro; IPR004005; Calici_coat.		
DR	InterPro; IPR008975; Viral_cap_coat.		
DR	Pfam; PF00915; Calici_coat; 1.		
SQ	SEQUENCE 550 AA; 60321 MW; 543265E29E05D1F0 CRC64;		
Query Match	100.0%;	Score 2906;	DB 12; Length 550;
Best Local Similarity	100.0%;	Pred. No. 1.3e-227;	
Matches 550;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Qy	1	MKMASNDAAPSNDGAANLVP	ANDVMALEPVVGGASIAAPVVGQNIIDPWIRENFVQAP 60
Db	1	MKMASNDAAPSNDGAANLVP	ANDVMALEPVVGGASIAAPVVGQNIIDPWIRENFVQAP 60
Qy	61	QGEFTVSPRNSPGEMLLNL	LELGPENLYSLSRMYNGYAGQMVGQVVLGNAGNFTAGKII 120

Wed Jun 2 09:13:25 2004

```

DR Pfm; PF00915; Calici coat; 1. 4B8AE1352B79D995 CRC64;
SQ SEQUENCE 550 AA; 60374 MW; 1; Mismatches 6; Indels 0; Gaps 0;

Query Match 98.5%; Score 2861; DB 12; Length 550;
Best Local Similarity 98.7%; Pred. NO. 6.1e-224;
Matches 543; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 1 MKWASNDAAAPSNDGAANLVPEANDEVMALEPVVVGASIAAPVVGQONIIDPWIRENFVQAP 60
DB 1 MKWASNDAAAPSNDGAANLVPEANDEVMALEPVVVGASIAAPVVGQONIIDPWIRENFVQAP 60
QY 61 QGFTVSPRNSPGEMLLNLELGPENLYLHLSRMVYAGGMQVQVVLGNAFTAGKII 120
DB 61 QGFTVSPRNSPGEMLLNLELGPENLYLHLSRMVYAGGMQVQVVLGNAFTAGKII 120
QY 121 FAAVPHFPVENISAAQITMCPHVIVDVRQLEPVLLPLDIRNRFHYNQENTPRMRLVA 180
DB 121 FAAVPHFPVENISAAQITMCPHVIVDVRQLEPVLLPLDIRNRFHYNQENTPRMRLVA 180
QY 181 MLYTPLRANGSDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTPFTLPIITLGLNSR 240
DB 181 MLYTPLRANGSDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTPFTLPIITLGLNSR 240
QY 241 FPAADMLYTDPNESIVVQPNQRCCTLDGTLQGTTLQVPTQICAPRGTLISQTARAADST 300
DB 241 FPAADMLYTDPNESIVVQPNQRCCTLDGTLQGTTLQVPTQICAPRGTLISQTARAADST 300
QY 301 DSPQARNHPLHVQVKNLDGTOYDPTDDIPAVLGAIDFKGTVFGVASQORDVSGQEQGHY 360
DB 301 DSPQARNHPLHVQVKNLDGTOYDPTDDIPAVLGAIDFKGTVFGVASQORDVSGQEQGHY 360
QY 361 ATRAHEAHIDTTPKYAPKLTILIKSGSDDFNTNQPIRFTPVGMGDNWRQWELPDYSG 420
DB 361 ATRAHEAHIDTTPKYAPKLTILIKSGSDDFNTNQPIRFTPVGMGDNWRQWELPDYSG 420
QY 421 RLTLNMNLAPAVSPSPGERILFFRSIVPSAGGSGYIDCLIPQEWVQHFYQEAAPSQS 480
DB 421 RLTLNMNLAPAVSPSPGERILFFRSIVPSAGGSGYIDCLIPQEWVQHFYQEAAPSQS 480
QY 481 AVALRVYVNPDTGRNIFPEAKLHREGFLTVANCGNPIVPPNGYFRFEAGNQFYTLAPM 540
DB 481 AVALRVYVNPDTGRNIFPEAKLHREGFLTVANCGNPIVPPNGYFRFEAGNQFYTLAPM 540
QY 541 GSGQGRRAQ 550
DB 541 GSGQGRRAQ 550

RESULT 3
Q918B3 PRELIMINARY; PRT; 550 AA.
AC Q918B3;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Baltimore/274/1993/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171840;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Baltimore/274/1993/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Baltimore/274/1993/US;
RX MEDLINE=97193806; PubMed=9041391;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Baltimore/269/1993/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414407; AAL12974.1; -
DR InterPro; IPR004005; Calici coat.
DR InterPro; IPR008975; Viral_Gap_coat.

```



RX MEDLINE=20266071; PubMed=10804147;  
RA Ando T., Noel J.S., Fankhauser R.L.;  
RT "Genetic classification of 'Norwalk-like viruses.'";  
RL J. Infect. Dis. 181:8336-8348(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=HU/NLV/Baltimore/274/1993/US;  
RA Ando T., Seo Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF414408; AAL12977.1; -  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 550 AA; 60300 MW; 843401DE6FFBC055 CRC64;  
  
Query Match 98.2%; Score 2855; DB 12; Length 550;  
Best Local Similarity 98.5%; Pred. No. 1.9e-223;  
Matches 542; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
QY 1 MKMASNDAAAPSDGAAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60  
DB 1 MKMASNDAAAPSDGAAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60  
  
QY 61 QGEFTVSPNSPCEMLLNLELGPENLPYLSHLSRMVYAGGVQVVLGNAFTAGKII 120  
DB 61 QGEFTVSPNSPCEMLLNLELGPENLPYLSHLSRMVYAGGVQVVLGNAFTAGKII 120  
  
QY 121 FFAAIDMLYTDNPNESIVVQPNQGRCTLDGTTQTLVPTQICAFRGTLISQTARAADST 300  
DB 121 FFAAIDMLYTDNPNESIVVQPNQGRCTLDGTTQTLVPTQICAFRGTLISQTARAADST 300  
  
QY 301 DSPQARARNHPLHVQVKNLGGTQYDPTDDIPAVLGAIDFKGTGFGVASQRDVSGQOQGHY 360  
DB 301 DSPQARARNHPLHVQVKNLGGTQYDPTDDIPAVLGAIDFKGTGFGVASQRDVSGQOQGHY 360  
  
QY 361 ATRAHAHIDTDPKYPKLGTLIKSGDDFNFTNQPIRFTPVGMGDNWNRQWELPDYSG 420  
DB 361 ATRAHAHIDTDPKYPKLGTLIKSGDDFNFTNQPIRFTPVGMGDNWNRQWELPDYSG 420  
  
QY 421 RLTLNNMLAPAVSPSPGGERILFFRSIVPSAGYSGYIDCLIPQEWVQHFYQEAAPQS 480  
DB 421 RLTLNNMLAPAVSPSPGGERILFFRSIVPSAGYSGYIDCLIPQEWVQHFYQEAAPQS 480  
  
QY 481 AVALVRYVNPDTGRNIFEAKLHREGFLTVANGCNPNPIVPPNGYFRFEAWNQFYTLAPM 540  
DB 481 AVALVRYVNPDTGRNIFEAKLHREGFLTVANGCNPNPIVPPNGYFRFEAWNQFYTLAPM 540  
  
QY 541 GSGQGRRAQ 550  
DB 541 GSGQGRRAQ 550

RESULT 4  
Q9IV37 PRELIMINARY; PRT; 550 AA.  
AC Q9IV37;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus HU/NLV/Seacroft/90/UK.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=122924;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN=HU/NLV/Seacroft/90/UK;  
RX MEDLINE=20404883; PubMed=10949950;  
RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,  
RA Clegg J.C., Chamberlain J., Brown D.W.G.;  
RT "Capsid protein diversity among 'Norwalk-like' viruses.";  
RL Virus Genes 20:227-236(2000).  
DR EMBL; AJ277620; CAB89101.1; -  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 550 AA; 60355 MW; 71AF0C5CA1DA0506 CRC64;  
  
Query Match 98.0%; Score 2847; DB 12; Length 550;  
Best Local Similarity 98.4%; Pred. No. 8.4e-223;  
Matches 541; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
  
QY 1 MKMASNDAAAPSDGAAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60  
DB 1 MKMASNDAAAPSDGAAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60  
  
QY 61 QGEFTVSPNSPCEMLLNLELGPENLPYLSHLSRMVYAGGVQVVLGNAFTAGKII 120  
DB 61 QGEFTVSPNSPCEMLLNLELGPENLPYLSHLSRMVYAGGVQVVLGNAFTAGKII 120  
  
QY 121 FFAAIDMLYTDNPNESIVVQPNQGRCTLDGTTQTLVPTQICAFRGTLISQTARAADST 300  
DB 121 FFAAIDMLYTDNPNESIVVQPNQGRCTLDGTTQTLVPTQICAFRGTLISQTARAADST 300  
  
QY 301 DSPQARARNHPLHVQVKNLGGTQYDPTDDIPAVLGAIDFKGTGFGVASQRDVSGQOQGHY 360  
DB 301 DSPQARARNHPLHVQVKNLGGTQYDPTDDIPAVLGAIDFKGTGFGVASQRDVSGQOQGHY 360  
  
QY 361 ATRAHAHIDTDPKYPKLGTLIKSGDDFNFTNQPIRFTPVGMGDNWNRQWELPDYSG 420  
DB 361 ATRAHAHIDTDPKYPKLGTLIKSGDDFNFTNQPIRFTPVGMGDNWNRQWELPDYSG 420  
  
QY 421 RLTLNNMLAPAVSPSPGGERILFFRSIVPSAGYSGYIDCLIPQEWVQHFYQEAAPQS 480  
DB 421 RLTLNNMLAPAVSPSPGGERILFFRSIVPSAGYSGYIDCLIPQEWVQHFYQEAAPQS 480  
  
QY 481 AVALVRYVNPDTGRNIFEAKLHREGFLTVANGCNPNPIVPPNGYFRFEAWNQFYTLAPM 540  
DB 481 AVALVRYVNPDTGRNIFEAKLHREGFLTVANGCNPNPIVPPNGYFRFEAWNQFYTLAPM 540  
  
QY 541 GSGQGRRAQ 550  
DB 541 GSGQGRRAQ 550

RESULT 5  
Q8V778 PRELIMINARY; PRT; 550 AA.  
AC Q8V778;  
DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE ORF2 protein (Capsid).  
OS Norwalk-like virus.  
OS Norwalk-like virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=11983, 95340;  
RN [1]  
RP SEQUENCE FROM N.A.

Wed Jun 2 09:13:25 2004

```

RC SPECIES=Norwalk virus; STRAIN=U4GII;
RA Kojima S., Kageyama T., Fukushi S., Hoshino F.B., Shinohara M.,
RA Uchida K., Natori K., Takeda N., Katayama K.;
RT "Genogroup-specific primers for detect Norwalk like virus.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama U4;
RA Katayama K.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama U4;
RX MEDLINE=22192455; PubMed=1202225;
RA Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,
RA Gojobori T., Takeda N.;
RT "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
RT Viruses.";
RL Virology 299:225-239 (2002).
DR EMBL; AB067538; BAB84143.1; -
DR EMBL; AB039777; BAC11822.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 550 AA; 60227 MW; 087973F2AAC5E789 CRC64;

Query Match 97.3%; Score 2828; DB 12; Length 550;
Best Local Similarity 97.8%; Pred. No. 3e-221;
Matches 538; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Qy 1 MKMASNDAAAPSDGAAANLVPEANDEVMALEPVVGASIAAPVVGQNIIDPWIRENFVQAP 60
Db 1 MKMASNDAAAPSDGAAANLVPEANDEVMALEPVVGASIAAPVVGQNIIDPWIRENFVQAP 60

Qy 61 QGEFTVSPRNSPGEMLNLELGPENLVLSRWYNGYAGGMQVQVVLGNAFTAGKII 120
Db 61 QGEFTVSPRNSPGEMLNLELGPENLVLSRWYNGYAGGMQVQVVLGNAFTAGKII 120

Qy 61 QGEFTVSPRNSPGEMLNLELGPENLVLSRWYNGYAGGMQVQVVLGNAFTAGKII 120
Db 61 QGEFTVSPRNSPGEMLNLELGPENLVLSRWYNGYAGGMQVQVVLGNAFTAGKII 120

Qy 121 FFAVPPHPPVENISAAQITMCPHIVDVRLQLEPVLPLDINRPFHYNQENTPRMLVA 180
Db 121 FFAVPPHPPVENISAAQITMCPHIVDVRLQLEPVLPLDINRPFHYNQENTPRMLVA 180

Qy 121 FFAVPPHPPVENISAAQITMCPHIVDVRLQLEPVLPLDINRPFHYNQENTPRMLVA 180
Db 121 FFAVPPHPPVENISAAQITMCPHIVDVRLQLEPVLPLDINRPFHYNQENTPRMLVA 180

Qy 181 MLYTFLRANGSGDVFVSCRVLTREPAPDFEFLVPPTVESKTKPFTLPILTGLGELNSR 240
Db 181 MLYTFLRANGSGDVFVSCRVLTREPAPDFEFLVPPTVESKTKPFTLPILTGLGELNSR 240

Qy 241 FFAAIDMLYTDPNESIVVQPNQGRCTLDGTQGTTLQVPTQICAPRGTLISQTARAADST 300
Db 241 FFAAIDMLYTDPNESIVVQPNQGRCTLDGTQGTTLQVPTQICAPRGTLISQTARAADST 300

Qy 301 DSPQARARNHPLHVQVKNLGGTQYDPTDDIPAVLGAIDFKGTGTFVGASQSDVSGQEQGHY 360
Db 301 DSPQARARNHPLHVQVKNLGGTQYDPTDDIPAVLGAIDFKGTGTFVGASQSDVSGQEQGHY 360

Qy 361 ATRAHEAHIDTDPKYPAPKLGITLILKSGSDDTNTQPIRFTPVGMGNNRWQELPDYSG 420
Db 361 ATRAHEAHIDTDPKYPAPKLGITLILKSGSDDTNTQPIRFTPVGMGNNRWQELPDYSG 420

Qy 541 GSGGRRRAQ 550
Db 541 GSGGRRRAQ 550

RESULT 6

```

RESULT 6



```

DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 547 AA; 59916 MW; 5C4F7272208B62B6 CRC64;

Query Match 90.5%; Score 2630.5; DB 12; Length 547;
Best Local Similarity 89.8%; Pred. No. 3.4e-205;
Matches 494; Conservative 27; Mismatches 26; Indels 3; Gaps 1;

QY 1 MKMASNDAAPSNDGAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60
DB 1 MKMASNDAAPSNDGAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60
QY 61 QGEFTVSPNSGEMLLNLELGPENLPYLSHLSRMVYAGVQVVLGNAFTAGKII 120
DB 61 QGEFTVSPNSGEMLLNLELGPENLPYLSHLSRMVYAGVQVVLGNAFTAGKII 120
QY 61 QGEFTVSPNSGEMLLNLELGPENLPYLSHLSRMVYAGVQVVLGNAFTAGKII 120
DB 61 QGEFTVSPNSGEMLLNLELGPENLPYLSHLSRMVYAGVQVVLGNAFTAGKII 120
QY 121 FFAVPPHFFPVENISAAQITMCPHVIDVVRQLEPVLPLPDINRPFHYNQETPRMRLVA 180
DB 121 FFAVPPHFFPVENISAAQITMCPHVIDVVRQLEPVLPLPDINRPFHYNQETPRMRLVA 180
QY 181 MLYTPLRANGSDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPFTLPLTLGELSNSR 240
DB 181 MLYTPLRANGSDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPFTLPLTLGELSNSR 240
QY 241 FPAIDMLYTDPNESIVVQPNQGRCTLDGTLQGTTLQVPTQICAFRGTLISQARAADST 300
DB 241 FPAIDMLYTDPNESIVVQPNQGRCTLDGTLQGTTLQVPTQICAFRGTLISQARAADST 300
QY 301 DSPQARNHPLHVQKNLDGTQDPTDDIPAVLGAIDFKGTGFGVASQRDVSGQOEGHY 360
DB 301 DSPQARNHPLHVQKNLDGTQDPTDDIPAVLGAIDFKGTGFGVASQRDVSGQOEGHY 360
QY 361 ATRAHEAHIDTDPKYAPKLTILIKSGSDDFNTNPIRFTPVGMGNNRWQELPDYSG 420
DB 361 ATRAHEAHIDTDPKYAPKLTILIKSGSDDFNTNPIRFTPVGMGNNRWQELPDYSG 420
QY 358 ATRAHEVHIDTTPRYTPKLGSLMYSESSDFDDGQPTFTPIGMGADWHQWELPEYSG 417
DB 358 ATRAHEVHIDTTPRYTPKLGSLMYSESSDFDDGQPTFTPIGMGADWHQWELPEYSG 417
QY 421 RLTLNMLNAPVSPFGERILFFRSIVPSAGGYSYIDCLIPQEWVQHFYQEAAPSQS 480
DB 421 RLTLNMLNAPVSPFGERILFFRSIVPSAGGYSYIDCLIPQEWVQHFYQEAAPSQS 480
QY 418 HLTLMNMLAPAVAPAFGERILFFRSIVPSAGGYSYIDCLIPQEWVQHFYQEAAPSQS 477
DB 418 HLTLMNMLAPAVAPAFGERILFFRSIVPSAGGYSYIDCLIPQEWVQHFYQEAAPSQS 477
QY 481 AVALRVYNPDGTGRNIFEAKLHREGFLTVANGNNPIVVPNGYFRFEAWNGQFYTLAPM 540
DB 481 AVALRVYNPDGTGRNIFEAKLHREGFLTVANGNNPIVVPNGYFRFEAWNGQFYTLAPM 540
QY 478 AVALRVYNPDGTGRNIFEAKLHREGFLTVANGNNPIVVPNGYFRFEAWNGQFYTLAPM 537
DB 478 AVALRVYNPDGTGRNIFEAKLHREGFLTVANGNNPIVVPNGYFRFEAWNGQFYTLAPM 537
QY 541 GSGQGRRAQ 550
DB 541 GSGQGRRAQ 550
QY 538 GTGQGRRAQ 547
DB 538 GTGQGRRAQ 547

RESULT 9
Q8V773 PRELIMINARY; PRT; 547 AA.
AC Q8V773;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF2 protein (Capsid).
OS Norwalk virus, and
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983, 95340;
RN [1]
RC SPECIES=Norwalk virus; STRAIN=U17GII;
RA Kojima S., Kageyama T., Fukushima S., Hoshino F.B., Shinohara M.,
RA Uchida K., Natori K., Takeda N., Katayama K.;
RT "Genogroup-specific primers for detect Norwalk like virus.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP SPECIES=Norwalk-like virus; STRAIN=Saitama U17;
RC
```

OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae.  
 OX NCBI\_TaxID=11975;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TV24;  
 RX MEDLINE=94202329; PubMed=8151799;  
 RA Lew J.F., Petric M., Kapikian A.Z., Jiang X., Estes M.K., Green K.Y.;  
 RT "Identification of Minireovirus as a Norwalk-like virus in pediatric  
 patients with gastroenteritis";  
 RL J. Virol. 68:3391-3396(1994).  
 DR EMBL; U02030; AAA18930.1; -  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 548 AA; 60034 MW; CCBE4A180E85008B CRC64;  
 Query Match 73.3%; Score 2129; DB 12; Length 548;  
 Best Local Similarity 71.5%; Pred. No. 2.1e-164;  
 Matches 397; Conservative 72; Mismatches 74; Indels 12; Gaps 4;  
 QY 1 MKMASNDAAPSNDGAANLVPEANDEVMALEPVVGASIAAPVVGQNIIDPMIRENFVQAP 60  
 DB 1 MKMASNDAAPSNDGAACLVPEINNEAMALEPVAGSAIAAPLTGQNIIDPMIRNFVQAP 60  
 QY 61 QGEFTVSPRNSPGEMLLNLGPELNPYLHSLRMVYAGGMQVQVVLGNAGFTAGKII 120  
 DB 61 QGEFTVSPRNSPGEMLLNLGPELNPYLHSLRMVYAGGMQVQVVLGNAGFTAGKII 120  
 QY 121 FAAPVPPHPPVENISAAQITMCPHVIVDVQLEPVLPLPDRNRFFHYNQENTPMRLVA 180  
 DB 121 FAAPVPPHPPVENISAAQITMCPHVIVDVQLEPVLPLPDRNRFFHYNQENTPMRLVA 180  
 QY 181 MLYTPLRA-NSGEDVFTVSCRVLTRPADDFEFTFLVPPPTVESKTKPFTLPILITGELS 239  
 DB 181 MLYTPLRA-NSGEDVFTVSCRVLTRPADDFEFTFLVPPPTVESKTKPFTLPILITGELS 239  
 QY 240 RPPAIDMLYTDPNESIVVQPNQGRCTLDGTLGGTTLQVPTQICAFRGTLSQTRAAADS 299  
 DB 240 RPPAIDMLYTDPNESIVVQPNQGRCTLDGTLGGTTLQVPTQICAFRGTLSQTRAAADS 299  
 QY 241 RFPVPIDSLHTSPENIVVQCNQGRVTLTGELMTTQLPSQICAFRGTLSQTRAAADS 300  
 DB 241 RFPVPIDSLHTSPENIVVQCNQGRVTLTGELMTTQLPSQICAFRGTLSQTRAAADS 300  
 QY 300 TDSPP-QRARNHPLHVQKNLGGTQYDPTDDIPAVLGAIDFKGTGFGVAGSQRDVSQEQG 358  
 DB 300 TDSPP-QRARNHPLHVQKNLGGTQYDPTDDIPAVLGAIDFKGTGFGVAGSQRDVSQEQG 358  
 QY 301 ADTPTPLRNFYHWHIQLDNLNGTTPYDPAEDIPAPLGTDFRGKVGVSQRNPDS 355  
 DB 301 ADTPTPLRNFYHWHIQLDNLNGTTPYDPAEDIPAPLGTDFRGKVGVSQRNPDS 355  
 QY 359 HYATRAEAHIDTTPKYPAPKLGTLIKSGSDDFNTNQPPIRFTPVGMGDNN---WRWEL 415  
 DB 359 HYATRAEAHIDTTPKYPAPKLGTLIKSGSDDFNTNQPPIRFTPVGMGDNN---WRWEL 415  
 QY 356 --TTTRAHEAKVDTTSGRTFPLKLSLEITTESDDFDNPTQSTKLPILITISEMNS 240  
 DB 356 --TTTRAHEAKVDTTSGRTFPLKLSLEITTESDDFDNPTQSTKLPILITISEMNS 240  
 QY 416 PDYSGRLLTNMNLAPAVSPSPGGERILFFRSIVPSAGYGGYDCLIPQEWVQHFYQEA 475  
 DB 416 PDYSGRLLTNMNLAPAVSPSPGGERILFFRSIVPSAGYGGYDCLIPQEWVQHFYQEA 475  
 QY 414 PNYSGQFTHMNLAPAVAPNPFGEQLLFFRSQSPSSGGRSNGVLDCLVPQEWVQHFYQES 473  
 DB 414 PNYSGQFTHMNLAPAVAPNPFGEQLLFFRSQSPSSGGRSNGVLDCLVPQEWVQHFYQES 473  
 QY 476 APSQSAVALRVYVNPDTGRNIFPEAKLHREGFLTVANCNNPITVPPNGYFRFESWVNPFF 535  
 DB 476 APSQSAVALRVYVNPDTGRNIFPEAKLHREGFLTVANCNNPITVPPNGYFRFESWVNPFF 535  
 QY 474 APAQTQVALRVYVNPDTGRVLFPEAKLHKGMTIAKNGSDSPITVPPNGYFRFESWVNPFF 533  
 DB 474 APAQTQVALRVYVNPDTGRVLFPEAKLHKGMTIAKNGSDSPITVPPNGYFRFESWVNPFF 533  
 QY 536 TLAPMGSCGQRRRAQ 550  
 DB 536 TLAPMGSCGQRRRAQ 550  
 QY 534 TLAPMGTCGRRRIQ 548  
 DB 534 TLAPMGTCGRRRIQ 548  
 RESULT 11  
 Q8V768  
 ID Q8V768 PRELIMINARY; PRT; 548 AA.  
 AC Q8V768  
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)  
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE ORF2 protein (Capsid).  
 OS Norwalk virus, and  
 OS Norwalk-like virus.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.

OX NCBI\_TaxID=11983, 95340;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Norwalk virus; STRAIN=U201GII;  
 RA Kojima S., Kageyama T., Fukushi S., Hoshino F.B., Shinohara M.,  
 RT Uchida K., Natori K., Takeda N., Katayama K.;  
 RL "Genogroup-specific primers for detect Norwalk like virus";  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Norwalk-like virus; STRAIN=Saitama U201;  
 RA Katayama K.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 [3]  
 RP SEQUENCE FROM N.A.  
 RC SPECIES=Norwalk-like virus; STRAIN=Saitama U201;  
 RX MEDLINE=22192455; PubMed=12202225;  
 RA Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,  
 RA Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,  
 RA Gojobori T., Takeda N.;  
 RT "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like  
 Viruses";  
 RL Virology 299:225-239(2002).  
 DR EMBL; AB057542; BAB84155.1; -  
 DR EMBL; AB039782; BAC11837.1; -  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 548 AA; 59949 MW; 589FC65F14B80F64 CRC64;  
 Query Match 73.3%; Score 2129; DB 12; Length 548;  
 Best Local Similarity 71.5%; Pred. No. 2.1e-164;  
 Matches 397; Conservative 71; Mismatches 75; Indels 12; Gaps 4;  
 QY 1 MKMASNDAAPSNDGAANLVPEANDEVMALEPVVGASIAAPVVGQNIIDPMIRENFVQAP 60  
 DB 1 MKMASNDAAPSNDGAAGLVPEINNEAMALEPVAGAAIAAPLTGQNIIDPMIRNFVQAP 60  
 QY 61 QGEFTVSPRNSPGEMLLNLGPELNPYLHSLRMVYAGGMQVQVVLGNAGFTAGKII 120  
 DB 61 QGEFTVSPRNSPGEMLLNLGPELNPYLHSLRMVYAGGMQVQVVLGNAGFTAGKII 120  
 QY 121 FAAPVPPHPPVENISAAQITMCPHVIVDVQLEPVLPLPDRNRFFHYNQENTPMRLVA 180  
 DB 121 FAAPVPPHPPVENISAAQITMCPHVIVDVQLEPVLPLPDRNRFFHYNQENTPMRLVA 180  
 QY 181 MLYTPLRA-NSGEDVFTVSCRVLTRPADDFEFTFLVPPPTVESKTKPFTLPILITGELS 239  
 DB 181 MLYTPLRA-NSGEDVFTVSCRVLTRPADDFEFTFLVPPPTVESKTKPFTLPILITGELS 239  
 QY 240 RPPAIDMLYTDPNESIVVQPNQGRCTLDGTLGGTTLQVPTQICAFRGTLSQTRAAADS 299  
 DB 240 RPPAIDMLYTDPNESIVVQPNQGRCTLDGTLGGTTLQVPTQICAFRGTLSQTRAAADS 299  
 QY 241 RFPVPIDSLHTSPENIVVQCNQGRVTLTGELMTTQLPSQICAFRGTLSQTRAAADS 300  
 DB 241 RFPVPIDSLHTSPENIVVQCNQGRVTLTGELMTTQLPSQICAFRGTLSQTRAAADS 300  
 QY 300 TDSPP-QRARNHPLHVQKNLGGTQYDPTDDIPAVLGAIDFKGTGFGVAGSQRDVSQEQG 358  
 DB 300 TDSPP-QRARNHPLHVQKNLGGTQYDPTDDIPAVLGAIDFKGTGFGVAGSQRDVSQEQG 358  
 QY 301 ADTPTPLRNFYHWHIQLDNLNGTTPYDPAEDIPAPLGTDFRGKVGVSQRNPDS 355  
 DB 301 ADTPTPLRNFYHWHIQLDNLNGTTPYDPAEDIPAPLGTDFRGKVGVSQRNPDS 355  
 QY 359 HYATRAEAHIDTTPKYPAPKLGTLIKSGSDDFNTNQPPIRFTPVGMGDNN---WRWEL 415  
 DB 359 HYATRAEAHIDTTPKYPAPKLGTLIKSGSDDFNTNQPPIRFTPVGMGDNN---WRWEL 415  
 QY 356 --TTTRAHEAKVDTTSGRTFPLKLSLEITTESDDFDNPTQSTKLPILITISEMNS 240  
 DB 356 --TTTRAHEAKVDTTSGRTFPLKLSLEITTESDDFDNPTQSTKLPILITISEMNS 240  
 QY 416 PDYSGRLLTNMNLAPAVSPSPGGERILFFRSIVPSAGYGGYDCLIPQEWVQHFYQEA 475  
 DB 416 PDYSGRLLTNMNLAPAVSPSPGGERILFFRSIVPSAGYGGYDCLIPQEWVQHFYQEA 475  
 QY 414 PNYSGQFTHMNLAPAVAPNPFGEQLLFFRSQSPSSGGRSNGVLDCLVPQEWVQHFYQES 473  
 DB 414 PNYSGQFTHMNLAPAVAPNPFGEQLLFFRSQSPSSGGRSNGVLDCLVPQEWVQHFYQES 473  
 QY 476 APSQSAVALRVYVNPDTGRNIFPEAKLHREGFLTVANCNNPITVPPNGYFRFESWVNPFF 535  
 DB 476 APSQSAVALRVYVNPDTGRNIFPEAKLHREGFLTVANCNNPITVPPNGYFRFESWVNPFF 535  
 QY 474 APAQTQVALRVYVNPDTGRVLFPEAKLHKGMTIAKNGSDSPITVPPNGYFRFESWVNPFF 533  
 DB 474 APAQTQVALRVYVNPDTGRVLFPEAKLHKGMTIAKNGSDSPITVPPNGYFRFESWVNPFF 533  
 QY 536 TLAPMGSCGQRRRAQ 550  
 DB 536 TLAPMGSCGQRRRAQ 550  
 QY 534 TLAPMGTCGRRRIQ 548  
 DB 534 TLAPMGTCGRRRIQ 548

```
RESULT 12
Q8V0P5 PRELIMINARY; PRT; 548 AA.
ID Q8V0P5;
AC Q8V0P5;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus Hu/NLV/GII/MD101-2/1987/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=159311;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/GII/MD101-2/1987/US;
RX MEDLINE=21666333; PubMed=11807686;
RA Green K.Y., Belliot G., Taylor J.L., Valdesuso J., Lew J.F.,
RA Kapiikian A.Z., Lin F.Y.C.;
RA "A Predominant Role for Norwalk-like Viruses as Agents of Epidemic
RT Gastroenteritis in Maryland Nursing Homes for the Elderly.";
RL J. Infect. Dis. 185:133-146(2002).
DR EMBL; AY030312; AAK54360.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR PFam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 53698 MW; 776F74B4E58B6CC7 CRC64;

Query Match 73.2%; Score 2126; DB 12; Length 548;
Best Local Similarity 71.5%; Pred. No. 3.6e-164;
Matches 397; Conservative 68; Mismatches 78; Indels 12; Gaps 4;

Qy 1 MKMASNDAAPSNDGAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAFLTCQNIIDPWIMNFVQAP 60
Qy 61 QGEFTVSPRSPGEMLNLELGPENLYLSHLSRMVNYAGAGVQVVLGNAFTAGKII 120
Db 61 QGEFTVSPRSPGEMLNLELGPENLYLSHLSRMVNYAGAGVQVVLGNAFTAGKVI 120
Qy 61 QGEFTVSPRSPGEMLNLELGPENLYLSHLSRMVNYAGAGVQVVLGNAFTAGKII 120
Db 61 QGEFTVSPRSPGEMLNLELGPENLYLSHLSRMVNYAGAGVQVVLGNAFTAGKVI 120
Qy 121 FFAVPPHPVENISAAQITMCPHVIVDVRQLEPVLPLDIRNFHYNGENTPRMLVA 180
Db 121 FFAVPPHPVENISAAQITMCPHVIVDVRQLEPVLPLDIRNFHYNGENTPRMLVA 180
Qy 121 FFAVPPHPVENISAAQITMCPHVIVDVRQLEPVLPLDIRNFHYNGENTPRMLVA 180
Db 121 FFAVPPHPVENISAAQITMCPHVIVDVRQLEPVLPLDIRNFHYNGENTPRMLVA 180
Qy 181 MLYTPLRA--NGEDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPLITISEMS 239
Db 181 MLYTPLRA--NGEDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPLITISEMS 239
Qy 181 MLYTPLRANSGDDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPLITISEMS 240
Db 181 MLYTPLRANSGDDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPLITISEMS 240
Qy 240 RPPAIDMLYTDPNESIYVQNGRCCTLDGTLQGTQTLVPTQICAFRGTLISQARAADS 299
Db 240 RPPAIDMLYTDPNESIYVQNGRCCTLDGTLQGTQTLVPTQICAFRGTLISQARAADS 299
Qy 241 RFPVPIDSLTSPTESIYVQNGRCCTLDGTLQGTQTLVPTQICAFRGTLISQARAADS 300
Db 241 RFPVPIDSLTSPTESIYVQNGRCCTLDGTLQGTQTLVPTQICAFRGTLISQARAADS 300
Qy 300 TD-SFQARNHPLHVQKNLDTQVDDTDDIPAVLGADFKGTGFGVVASQRDVSGQEQ 358
Db 300 TD-SFQARNHPLHVQKNLDTQVDDTDDIPAVLGADFKGTGFGVVASQRDVSGQEQ 358
Qy 301 ADTATPRLFNHHWHIQLDNLNGTTPYDPAEDIPAPLGTPDFRGKVGFGVVASQRPDS 355
Db 301 ADTATPRLFNHHWHIQLDNLNGTTPYDPAEDIPAPLGTPDFRGKVGFGVVASQRPDS 355
Qy 359 HYATRAHEAHIDTTPKVPKLGTLIKSGDDFNTPQIRFTPVGMGDN--WRQWEL 415
Db 359 HYATRAHEAHIDTTPKVPKLGTLIKSGDDFNTPQIRFTPVGMGDN--WRQWEL 415
Qy 356 --TTAHEAKVDTTSGRFTPKLGSLETTESDDFPNQSTKFTPVGIGVDNEADFQW 413
Db 356 --TTAHEAKVDTTSGRFTPKLGSLETTESDDFPNQSTKFTPVGIGVDNEADFQW 413
Qy 416 PDYSGELTNMNLAPAVSPFGERILFPRSVPSAGGVSIGYIDCLIPQWVQHFYQEA 475
Db 416 PDYSGELTNMNLAPAVSPFGERILFPRSVPSAGGVSIGYIDCLIPQWVQHFYQEA 475
Qy 414 PDYSGOFTNMNLAPAVSPFGERILFPRSVPSAGGVSIGYIDCLIPQWVQHFYQES 473
Db 414 PDYSGOFTNMNLAPAVSPFGERILFPRSVPSAGGVSIGYIDCLIPQWVQHFYQES 473
Qy 476 APSQSAVALRVYNPDTGRNIPEAKLHREGFLTVANCNNPIVPPNGYFRFEAGNQFY 535
Db 476 APSQSAVALRVYNPDTGRNIPEAKLHREGFLTVANCNNPIVPPNGYFRFEAGNQFY 535
Qy 474 APAQTQVALRVYNPDTGRVLFKALHKLGFMTIAKNGDSPITVPPNGYFRFESWVNP 533
Db 474 APAQTQVALRVYNPDTGRVLFKALHKLGFMTIAKNGDSPITVPPNGYFRFESWVNP 533
Qy 536 TLAPMGSGQGRRAQ 550
Db 536 TLAPMGSGQGRRAQ 550
Qy 534 TLAPMGSGQGRRIQ 548
Db 534 TLAPMGSGQGRRIQ 548

RESULT 13
Q8V0P4 PRELIMINARY; PRT; 548 AA.
ID Q8V0P4;
AC Q8V0P4;
DT 01-MAR-2002 (TREMELrel. 20, Created)
DT 01-MAR-2002 (TREMELrel. 20, Last sequence update)
DT 01-OCT-2003 (TREMELrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus Hu/NLV/GII/MD134-10/1987/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=159312;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/GII/MD134-10/1987/US;
RX MEDLINE=21666333; PubMed=11807686;
RA Green K.Y., Belliot G., Taylor J.L., Valdesuso J., Lew J.F.,
RA Kapiikian A.Z., Lin F.Y.C.;
RA "A Predominant Role for Norwalk-like Viruses as Agents of Epidemic
RT Gastroenteritis in Maryland Nursing Homes for the Elderly.";
RL J. Infect. Dis. 185:133-146(2002).
DR EMBL; AY030313; AAK54361.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
DR PFam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 53899 MW; 0C02C5D953500351 CRC64;

Query Match 73.0%; Score 2121; DB 12; Length 548;
Best Local Similarity 71.4%; Pred. No. 9.2e-164;
Matches 396; Conservative 69; Mismatches 78; Indels 12; Gaps 4;

Qy 1 MKMASNDAAPSNDGAANLVPEANDEVMALEPVVVGASIAAPVVGQNIIDPWIRENFVQAP 60
Db 1 MKMASNDAAPSNDGAAGLVPEINNEAMALDPVAGAAIAAFLTCQNIIDPWIMNFVQAP 60
Qy 61 QGEFTVSPRSPGEMLNLELGPENLYLSHLSRMVNYAGAGVQVVLGNAFTAGKII 120
Db 61 QGEFTVSPRSPGEMLNLELGPENLYLSHLSRMVNYAGAGVQVVLGNAFTAGKVI 120
Qy 61 QGEFTVSPRSPGEMLNLELGPENLYLSHLSRMVNYAGAGVQVVLGNAFTAGKII 120
Db 61 QGEFTVSPRSPGEMLNLELGPENLYLSHLSRMVNYAGAGVQVVLGNAFTAGKVI 120
Qy 121 FFAVPPHPVENISAAQITMCPHVIVDVRQLEPVLPLDIRNFHYNGENTPRMLVA 180
Db 121 FFAVPPHPVENISAAQITMCPHVIVDVRQLEPVLPLDIRNFHYNGENTPRMLVA 180
Qy 121 FFAVPPHPVENISAAQITMCPHVIVDVRQLEPVLPLDIRNFHYNGENTPRMLVA 180
Db 121 FFAVPPHPVENISAAQITMCPHVIVDVRQLEPVLPLDIRNFHYNGENTPRMLVA 180
Qy 181 MLYTPLRA--NGEDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPLITISEMS 239
Db 181 MLYTPLRA--NGEDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPLITISEMS 239
Qy 181 MLYTPLRANSGDDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPLITISEMS 240
Db 181 MLYTPLRANSGDDVFTVSCRVLTRPAPDFEFTFLVPPTVESKTKPTLPLITISEMS 240
Qy 240 RPPAIDMLYTDPNESIYVQNGRCCTLDGTLQGTQTLVPTQICAFRGTLISQARAADS 299
Db 240 RPPAIDMLYTDPNESIYVQNGRCCTLDGTLQGTQTLVPTQICAFRGTLISQARAADS 299
Qy 241 RFPVPIDSLTSPTESIYVQNGRCCTLDGTLQGTQTLVPTQICAFRGTLISQARAADS 300
Db 241 RFPVPIDSLTSPTESIYVQNGRCCTLDGTLQGTQTLVPTQICAFRGTLISQARAADS 300
Qy 300 TD-SFQARNHPLHVQKNLDTQVDDTDDIPAVLGADFKGTGFGVVASQRDVSGQEQ 358
Db 300 TD-SFQARNHPLHVQKNLDTQVDDTDDIPAVLGADFKGTGFGVVASQRDVSGQEQ 358
Qy 301 ADTATPRLFNHHWHIQLDNLNGTTPYDPAEDIPAPLGTPDFRGKVGFGVVASQRPDS 355
Db 301 ADTATPRLFNHHWHIQLDNLNGTTPYDPAEDIPAPLGTPDFRGKVGFGVVASQRPDS 355
Qy 359 HYATRAHEAHIDTTPKVPKLGTLIKSGDDFNTPQIRFTPVGMGDN--WRQWEL 415
Db 359 HYATRAHEAHIDTTPKVPKLGTLIKSGDDFNTPQIRFTPVGMGDN--WRQWEL 415
Qy 356 --TTAHEAKVDTTSGRFTPKLGSLETTESDDFPNQSTKFTPVGIGVDNEADFQW 413
Db 356 --TTAHEAKVDTTSGRFTPKLGSLETTESDDFPNQSTKFTPVGIGVDNEADFQW 413
Qy 416 PDYSGELTNMNLAPAVSPFGERILFPRSVPSAGGVSIGYIDCLIPQWVQHFYQEA 475
Db 416 PDYSGELTNMNLAPAVSPFGERILFPRSVPSAGGVSIGYIDCLIPQWVQHFYQEA 475
Qy 414 PDYSGOFTNMNLAPAVSPFGERILFPRSVPSAGGVSIGYIDCLIPQWVQHFYQES 473
Db 414 PDYSGOFTNMNLAPAVSPFGERILFPRSVPSAGGVSIGYIDCLIPQWVQHFYQES 473
Qy 476 APSQSAVALRVYNPDTGRNIPEAKLHREGFLTVANCNNPIVPPNGYFRFEAGNQFY 535
Db 476 APSQSAVALRVYNPDTGRNIPEAKLHREGFLTVANCNNPIVPPNGYFRFEAGNQFY 535
Qy 474 APAQTQVALRVYNPDTGRVLFKALHKLGFMTIAKNGDSPITVPPNGYFRFESWVNP 533
Db 474 APAQTQVALRVYNPDTGRVLFKALHKLGFMTIAKNGDSPITVPPNGYFRFESWVNP 533
Qy 536 TLAPMGSGQGRRAQ 550
Db 536 TLAPMGSGQGRRAQ 550
Qy 534 TLAPMGSGQGRRIQ 548
Db 534 TLAPMGSGQGRRIQ 548
```

```

RESULT 14
ID Q91722 PRELIMINARY; PRT; 548 AA.
AC Q91722;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Brattleboro/321/1995/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171847;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;
RX MEDLINE=97193806; PubMed=94041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
characterized small round-structured viruses involved in outbreaks of
nonbacterial acute gastroenteritis in the United States, 1990 to
1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Brattleboro/321/1995/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414415; AAL12998.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59933 MW; 7B8AFF9AF1469158 CRC64;

Query Match 72.8%; Score 2117; DB 12; Length 548;
Best Local Similarity 71.2%; Pred. No. 1.9e-163;
Matches 395; Conservative 71; Mismatches 77; Indels 12; Gaps 4;

QY 1 MKWASNDAAAPSNDGAANLVPEANDEVMALEPVVGSIAAPVVGQNNIIDPWIRENFVQAP 60
Db 1 MKWASNDAAAPSNDGAAGLVPEINNEAMALEPVAGAIAAPLTGQQNIIDPWIMNWFVQAP 60
QY 61 QGEFTVSPNSGEMLLNLELGPENLPYLHSLRMVYAGQVQVVLGNAFTAGKII 120
Db 61 QGEFTVSPNSGEMLLNLELGPENLPYLHSLRMVYAGQVQVVLGNAFTAGKII 120
QY 121 FAAVPHPPVENISAAQITMCQPHVIVDVRQLPVLPLDINRPFHYNQENTPRMLVA 180
Db 121 FAAIPNFPIDNLAAQITMCQPHVIVDVRQLPVLPLDINRPFHYNQENTPRMLVA 180
QY 181 MLYTPLRA-NSGEDVFTVSCRVLTAPDFEFTFLVPVTSKTPFTLPILTLAGELSNS 239
Db 181 MLYTPLRANSGDDVFTVSCRVLTAPDFEFTFLVPVTSKTPFTLPILTLAGELSNS 240
QY 240 RPPAIDMLYTPDPSIVVQPNQGRCTLDGTLOGTTLVPTQICAFRGTLTRSTASDQ 299
Db 240 RPPAIDMLYTPDPSIVVQPNQGRCTLDGTLOGTTLVPTQICAFRGTLTRSTASDQ 299

```

```

Db 241 RPPVIDSLHTSPTEINIVQONGRVTLDGELMGTTLQLPSQICAFRGTLTRSTASDQ 300
QY 300 TDSP-QRARNHPLHVQVKNLQDGTQYDDIPAVLGAIDFKGTVFGVASQRDVSGQBOG 358
Db 301 ADTPPRLFNHRWHIQLONLNGTPYPAEDIPAPLGTDPFRGKVFVASORNPDS----- 355
QY 359 HYATRAHEAHIDTTPKYPAPKGLTTLIKSGSDDFNTNQPIRFTPVGMGDN---WROWEL 415
Db 356 --TTTRAHEAKVDTTSDRFTPKLSLEIITESGDFDNTQSTKFTPVGIGVDNEAEFQOWSL 413
QY 416 PDYSGRLTLNMLAPAVSPSPFGERILFFRSIVPSAGGYGSGYIDCLIPQEWVQHPYQBA 475
Db 414 PNYSCQFTNMLAPAVAPNPFGEQLLFFRSOLPSSGGRSNGVLCLIPQEWVQHPYQBS 473
QY 476 APSQSAVALRVYVNPDTGRTNIRPEAKLHREGFUTVANCNNPIVVPNGYFRFEAMGNOPY 535
Db 474 APAQTOVALRVYVNPDTGRTNIRPEAKLHREGFUTVANCNNPIVVPNGYFRFEAMGNOPY 533
QY 536 TLAPNGSGQGRRAQ 550
Db 534 TLAPNGTNGRRRIQ 548

RESULT 15
Q91725 PRELIMINARY; PRT; 548 AA.
ID Q91725;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE Capsid protein.
GN ORF2.
OS Norwalk-like virus NLV/Towson/313/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=171846;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Towson/313/1994/US;
RX MEDLINE=97193806; PubMed=94041391;
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577(1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/Towson/313/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
characterized small round-structured viruses involved in outbreaks of
nonbacterial acute gastroenteritis in the United States, 1990 to
1995.";
RL J. Med. Virol. 53:372-383(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Towson/313/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk-like viruses.'";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/Towson/313/1994/US;
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF414414; AAL12995.1; -
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59993 MW; 114F3907B3A26D89 CRC64;

```

Query Match	72.8%	Score 2117	DB 12	Length 548	
Best Local Similarity	71.4%	Pred. No. 1.9e-163			
Matches 396	Conservative 71	Mismatches 76	Indels 12	Gaps 4	
Qy	1	MKKASNDAA	SPNDGAANL	VPENDVMALPEPVVVGASIAAPVVGQOQNIIDPWIRENFVQAP 60	
Db	1	MKKASNDAA	SPNDGAAGL	VPENINNEAMALPFAVGAATAAPLTGQOQNIIDPWIMNNFVQAP 60	
Qy	61	QGEFTVSP	RNSPGCEMLNLEL	GPENLVLSHRMVTYAGGQVGVVLGNAFTAGKII 120	
Db	61	QGEFTVSP	RNSPGCEVLLNLEL	GPENLVLAHARMYAGGFEVGVVLGNAFTAGKVI 120	
Qy	121	FAAVPPHFP	VENISAAQITMC	PHVIVDVRQLEPVLPLDIRNRPFFHYNQENTPBMRLVA 180	
Db	121	FAAIPNFPI	DNLSAAQITMC	PHVIVDVRQLEPINLMPDVRNFPFFHYNQGDSRLRLIA 180	
Qy	181	MLYTPURA	-NSGEDVFTV	SCRVLTRPADPDEFTFLVPPTVESKTKFTPLITLAEGLNS 239	
Db	181	MLYTPURAN	SGDDVFTV	SCRVLTRPSPDFSFNLFVPPTVESKTKFTPLITLISEMNS 240	
Qy	240	RPFAADIM	LYTPDNESI	VVQPPQNGRCTLDTGLQTTQLVPTQICAPRGTLISOTARAADS 299	
Db	241	RPFPVP	IDSLSHTSP	TENIVVQCONGRVITLUGELMTGTLQLFSQICAFRGTLTRSTSEASDQ 300	
Qy	300	TDSP-OR	ARNHPLHVQ	KNLQDGTQDPTDDIPAVLGAI	DPKGTVFGVASORDVSGQOEG 358
Db	301	ADTFLR	LFNHRWHI	QLDNLNGTVPDAEDIPAPLGT	PDFRGKVFVGVASORPDS ----- 355
Qy	359	HYATRAE	AAHIDTTDP	KYAPKGLTILIKSGSD	DFNTQPIRFTVPMGMDNN---WRQWEL 415
Db	356	--TTTRA	EAKVDYTS	GRGFTPKLSGLIEIT	ESDDFTNQSTKFTFVGLVDNSEFEQOWSL 413
Qy	416	PYXSGRLT	LNMLNAP	AVSPSPFGERIL	FLFRSIVSAGYGSYGIDCLIPQEWQHFQEA 475
Db	414	PNYSGQ	FTHNNMLN	APAVNPLGEQLL	FFRSQPSGGRSNGVLDCILVPQEWQHFQES 473
Qy	476	APSQSA	VALVRYNP	PDGTGRNI	PEAKLHREGFLTVANCNNPITVPPNGYRFFFAWGNQFY 535
Db	474	APATQ	VALVRYNP	PDGTGRVL	PEAKLHLKGFMTIAKNGDSPLITVPPNGYRFFFSWNPFY 533
Qy	536	TLAPMG	SGQGRRAQ	550	
Db	534	TLAPMG	TNGRRRIQ	548	

Search completed: June 1, 2004, 13:53:23  
Job time : 34.0218 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:33:56 ; Search time 46.4752 Seconds  
(without alignments)  
3289.030 Million cell updates/sec

Title: US-09-926-799-11  
Perfect score: 2894  
Sequence: 1 MMSANDAPSSDGAAGLVP.....VNFYSLAPVGTGKRRRVQ 541

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

- Database : A\_Geneseq\_294an04.\*
- 1: Geneseqp1980s.\*
  - 2: Geneseqp1990s.\*
  - 3: Geneseqp2000s.\*
  - 4: Geneseqp2001s.\*
  - 5: Geneseqp2002s.\*
  - 6: Geneseqp2003as.\*
  - 7: Geneseqp2003bs.\*
  - 8: Geneseqp2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2894	100.0	541	AAB49710	Aab49710 Small rou
2	1678.5	58.0	548	AAB49705	Aab49705 Small rou
3	1656	57.2	535	AAB49707	Aab49707 Small rou
4	1650.5	57.0	540	AAB49706	Aab49706 Small rou
5	1640.5	56.7	548	AAB49709	Aab49709 Small rou
6	1613.5	55.8	550	AAB49703	Aab49703 Small rou
7	1569	54.2	539	AAB49704	Aab49704 Small rou
8	1538.5	53.2	542	AAB49708	Aab49708 Small rou
9	1013	35.0	545	AAB49700	Aab49700 Small rou
10	1008.5	34.8	530	AAB50972	Aab50972 Small rou
11	1008.5	34.8	530	ADC72176	Aab72176 Norwalk v
12	1005.5	34.7	530	AAB49701	Aab49701 Small rou
13	984.5	34.0	530	AAB57091	Aab57091 Small rou
14	970	33.5	544	AAB49703	Aab49703 Small rou
15	967	33.4	546	AAB49702	Aab49702 Small rou
16	309	10.7	579	AAB08143	Aab08143 RHDV caps
17	267.5	9.2	622	AAB47045	Aab47045 Feline ca
18	265.5	9.2	623	AAB47044	Aab47044 Feline ca
19	265.5	9.2	623	AAB47043	Aab47043 Feline ca
20	264	9.1	668	AAB67462	Aab67462 Amino aci
21	262.5	9.1	547	AAM50108	Aam50108 Feline ca
22	262.5	9.1	671	AAM50107	Aam50107 Feline ca
23	254	8.8	668	AAR10686	Aar10686 Feline ca
24	254	8.8	668	Aae04304	Aae04304 Feline ca
25	250.5	8.7	669	AAB67461	Aab67461 Amino aci

26	200	6.9	40	5	AAU91274	Aau91274 Norwalk v
27	174	6.0	40	5	AAU91273	Aau91273 Norwalk v
28	119.5	4.1	2209	1	AAP20037	Aap20037 Sequence
29	119	4.1	934	1	AAP20016	Aap20016 Sequence
30	118.5	4.1	1147	5	ABB76724	Abb76724 Foot and
31	117	4.0	2206	2	AAE22210	Aae22210 True type
32	109.5	3.8	1045	7	ADE72434	Ade72434 Human end
33	109.5	3.8	1070	7	ADE72427	Ade72427 Human end
34	109.5	3.8	1093	6	AAO23110	Aao23110 Lig-1 'hu
35	109.5	3.8	1093	7	ADE72425	Ade72425 Human end
36	109.5	3.8	1101	2	AAV08008	Aav08008 Human HLI
37	109.5	3.8	1196	7	ADY72433	Ade72433 Human end
38	108.5	3.7	861	6	ABU38084	Abu38084 Protein e
39	107	3.7	607	5	ABB53295	Abb53295 Human pol
40	106.5	3.7	767	4	AAM25696	Aam25696 Human pro
41	106.5	3.7	861	6	ABF77378	Abf77378 N. gonorr
42	106.5	3.7	861	6	ABU37406	Abu37406 Protein e
43	106.5	3.7	1070	5	ABG69674	Abg69674 Human sec
44	106.5	3.7	1077	5	ABP70116	Abp70116 Human NOV
45	106.5	3.7	1093	5	ABG69660	Abg69660 Human sec

ALIGNMENTS

RESULT 1  
AAB49710  
ID AAB49710 standard; protein; 541 AA.  
XX  
AC AAB49710;  
XX  
DT 04-APR-2001 (first entry)  
XX  
DE Small round structured virus protein SEQ ID 11.  
XX  
KW Small round structured virus; SRSV; food poisoning.  
XX  
OS Small round structured virus.  
XX  
PN WO200079280-A1.  
XX  
PD 28-DEC-2000.  
XX  
PF 22-JUN-2000; 2000WO-JP004095.  
XX  
PR 22-JUN-1999; 99JP-00175928.  
XX  
PA (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
PA  
Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
WPI: 2001-080848/09.  
N-PSDB; AAF29151.  
Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.  
Claim 1: Page 64-66; 84pp; Japanese.  
This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks  
SQ Sequence 541 AA;  
Query Match 100.0%; Score 2894; DB 4; Length 541;  
Best Local Similarity 100.0%; Pred. No. 5e-258;

Wed Jun 2 09:13:25 2004

Matches 541; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MKWASNDAPSDGAGLVEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVOAP 60  
 DB 1 MKWASNDAPSDGAGLVEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVOAP 60

QY 61 AGEFTVSPRNSPGEILLDLELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNATAGKII 120  
 DB 61 AGEFTVSPRNSPGEILLDLELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNATAGKII 120

QY 121 FPAIPGPFYENLSPQITMCPHVIIDVROLEPFLPMPDIWNNFFHYNQNDPKRLVA 180  
 DB 121 FPAIPGPFYENLSPQITMCPHVIIDVROLEPFLPMPDIWNNFFHYNQNDPKRLVA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPPVESKTKQFALPILKISEMTNS 240  
 DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPPVESKTKQFALPILKISEMTNS 240

QY 241 RFPVVDVMTARNENOVQVQNGRVTLTGELLGTTPLLAIVNICKFKGEVIAKNGDVRSY 300  
 DB 241 RFPVVDVMTARNENOVQVQNGRVTLTGELLGTTPLLAIVNICKFKGEVIAKNGDVRSY 300

QY 301 RMDMEITNTDGTIDPTDPTGPIGSPDFQILFGVASORNKNEQNPATRAHEAINTGG 360  
 DB 301 RMDMEITNTDGTIDPTDPTGPIGSPDFQILFGVASORNKNEQNPATRAHEAINTGG 360

QY 361 DHLCPQISSEIYLTSPNIIIRCTNPQLPQSGILRGITILRSNCHCHDMVGTSPPTTWP 420  
 DB 361 DHLCPQISSEIYLTSPNIIIRCTNPQLPQSGILRGITILRSNCHCHDMVGTSPPTTWP 420

QY 421 QQWRRCRSGNCCSGHRYFPVVMNRVTWIVLSHSGFSTSTRKLPQLNLRWPLIRFIN 480  
 DB 421 QQWRRCRSGNCCSGHRYFPVVMNRVTWIVLSHSGFSTSTRKLPQLNLRWPLIRFIN 480

QY 481 PDTGRVLFEARLHKGFTITVAHTGNDIVMPPNGYFPEAWNQFYSLAPVGTGKRRRV 540  
 DB 481 PDTGRVLFEARLHKGFTITVAHTGNDIVMPPNGYFPEAWNQFYSLAPVGTGKRRRV 540

QY 541 Q 541  
 DB 541 Q 541

RESULT 2  
 AAB49705  
 ID AAB49705 standard; protein; 548 AA.

XX AAB49705;  
 XX  
 XX 04-APR-2001 (first entry)  
 XX Small round structured virus protein SEQ ID 6.  
 XX Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.  
 XX W0200079280-A1.  
 XX 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.  
 XX 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
 XX (DENK-) DENKA SEIKEN KK.

XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
 XX WPI; 2001-080848/09.  
 XX N-PSDB; AAF29146.

PT Kit for the detection and typing of small round-structured virus (SRSV)  
 PT strains for investigation of food poisoning outbreaks, contains  
 PT antibodies.

XX Claim 1; Page 52-54; 84pp; Japanese.

XX This invention relates to a kit for the detection and typing of small  
 CC round structured virus (SRSV) strains. The kit contains antibodies  
 CC directed against peptides represented in sequences AAB49700 - AAB49710,  
 CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
 CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
 CC used for detecting and typing strains of SRSV in order to prevent the  
 CC spread of infection and to examine the epidemiology of outbreaks

XX Sequence 548 AA;

QY Query Match 58.0%; Score 1678.5; DB 4; Length 548;  
 DB Best Local Similarity 58.9%; Pred. No. 9.9e-146; Indels 45; Gaps 7;  
 DB Matches 334; Conservative 67; Mismatches 121;

QY 1 MKWASNDAPSDGAGLVEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVOAP 60  
 DB 1 MKWASNDAPSDGAGLVEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVOAP 60

QY 61 AGEFTVSPRNSPGEILLDLELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNATAGKII 120  
 DB 61 AGEFTVSPRNSPGEILLDLELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNATAGKII 120

QY 121 FPAIPGPFYENLSPQITMCPHVIIDVROLEPFLPMPDIWNNFFHYNQNDPKRLVA 180  
 DB 121 FPAIPGPFYENLSPQITMCPHVIIDVROLEPFLPMPDIWNNFFHYNQNDPKRLVA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPPVESKTKQFALPILKISEMTNS 240  
 DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPPVESKTKQFALPILKISEMTNS 240

QY 241 RFPVVDVMTARNENOVQVQNGRVTLTGELLGTTPLLAIVNICKFKGEVIAKNGDVRS- 299  
 DB 241 RFPVVDVMTARNENOVQVQNGRVTLTGELLGTTPLLAIVNICKFKGEVIAKNGDVRS- 299

QY 300 -----YRMDMEITNTDGTIDPTDPTGPIGSPDFQILFGVASORNKNEQNP 348  
 DB 300 QADTATPRLFNYYWHVQLDNLNGTYPDAEDIPGLGTDPFRGKVGVSQNRNLDJSDN---T 356

QY 349 TRAHEAINTGGDHLCPQISSEIYLTSPNIIIRCTNPQLPQSGILRGITILRSN----- 403  
 DB 349 TRAHEAINTGGDHLCPQISSEIYLTSPNIIIRCTNPQLPQSGILRGITILRSN----- 403

QY 357 TRAHEAKVDITTAGRFTPKLSLEISTDSDDDFDQ-NQPTKFTPVG-----IGVDNEAEFQ 409  
 DB 357 TRAHEAKVDITTAGRFTPKLSLEISTDSDDDFDQ-NQPTKFTPVG-----IGVDNEAEFQ 409

QY 404 -----GHCHDMVGTSPPTTPTWPOQWRRCRSGNCCSGHRYFPVVMNRVTWIVLS 454  
 DB 404 -----GHCHDMVGTSPPTTPTWPOQWRRCRSGNCCSGHRYFPVVMNRVTWIVLS 454

QY 410 QWSLDPDYSQFTHNNWNLAPAVAFNFFBQQLLFFRSQSPSGRSGNGLDCLVPOEWOHF 469  
 DB 410 QWSLDPDYSQFTHNNWNLAPAVAFNFFBQQLLFFRSQSPSGRSGNGLDCLVPOEWOHF 469

QY 455 HKSFGSTSTRKLPQLNLRWPLIRFINPDTGRVLFEARLHKGFTITVAHTGNDIVMPPNG 514  
 DB 455 HKSFGSTSTRKLPQLNLRWPLIRFINPDTGRVLFEARLHKGFTITVAHTGNDIVMPPNG 514

QY 470 YQESAPAQOV-----ALVRYVNPDTGKVLFEAKLHKLGFMTIANGDSPTVPNG 521  
 DB 470 YQESAPAQOV-----ALVRYVNPDTGKVLFEAKLHKLGFMTIANGDSPTVPNG 521

QY 515 YFRFEAWNQFYSLAPVGTGKRRVQ 541

DB 522 YFRFEAWNQFYSLAPVGTGKRRVQ 548

RESULT 3  
 AAB49707  
 ID AAB49707 standard; protein; 535 AA.

XX AAB49707;  
 XX

XX 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 8.

XX Small round structured virus; SRSV; food poisoning.

XX

OS Small round structured virus.  
PN WO200079280-A1.  
XX 28-DEC-2000.  
XX 22-JUN-2000; 2000WO-JP004095.  
XX 22-JUN-1999; 99JP-00175928.  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX WPI; 2001-080848/09.  
DR N-PSDB; AAF29148.  
XX Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
PS Claim 1; Page 57-59; 84pp; Japanese.  
XX This invention relates to a kit for the detection and typing of small  
CC round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX Sequence 535 AA;  
Query Match 57.2%; Score 1656; DB 4; Length 535;  
Best Local Similarity 59.2%; Pred. No. 1.1e-143;  
Matches 330; Conservative 67; Mismatches 122; Indels 38; Gaps 7;  
QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVQAP 60  
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVQAP 60  
QY 61 AGEFTVSPRNSGGEILLDELGPDLPNLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120  
DB 61 AGEFTVSPRNSGGEILLDELGPDLPNLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120  
QY 121 FAAIPPGPYENLSPSQITMCPHVIIDVROLEPFLPMPDIWNFFHYNQGNPKRLVA 180  
DB 121 FAAIPPGPYENLSPSQITMCPHVIIDVROLEPFLPMPDIWNFFHYNQGNPKRLVA 180  
QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNS 240  
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNS 240  
QY 241 RPPVPDVMYATARNENQVVPQNGRVLTDGELLGTTPLLAIVNICFKGEVIK-NGDVRS 299  
DB 241 RPPVPDVMYATARNENQVVPQNGRVLTDGELLGTTPLLAIVNICFKGEVIK-NGDVRS 299  
QY 300 YRMDIEITNTDCTPDIPTEDTPGIGSDPFGQILFVASQRNKNQNPATRAHEAINTG 359  
DB 299 HWNNQVNTINGTPDPTGDVPAFLGTDFSGKLFGLVLSQR---DHDNACRSHDAVIATN 355  
QY 360 GDHLCPOISSSIYLTSPNLIACITNPQPLPQSGL-----RGTLIRSDNGH 405  
DB 356 SAKFTPKLGAIGTWEEDDVHINQTKFTPVGLFENEGFNQWTLPNYSGALTL----- 409  
QY 406 CHDMVGTSTPTTWPQQWRRCRSGNCCSSGHRYFPVVMNRVTVIV-LSHKSGFSTSTR 464  
DB 410 --NMGLAPVAPTFFGEQILFRSHIPLKGGVADPVIDCLLPQEWIQHLYQESAPSQSD- 466  
QY 465 KLPQLNLRPLRFNPDGTGRVLPRLHKGQFIVTANGDNPVMPNGYFRFEAWNQ 524  
DB 467 -----VALIRFNPDTGRVLFPAKLHRSYITVANTGSRPIVVPANGYFRFDTWNQ 518

QY 525 FYSLAPVGTGKRRRVQ 541  
DB 519 FYSLAPMGTGNGRRVQ 535  
RESULT 4  
AAB49706  
ID AAB49706 standard; protein; 540 AA.  
XX AAB49706;  
XX 04-APR-2001 (first entry)  
XX Small round structured virus protein SEQ ID 7.  
XX Small round structured virus; SRSV; food poisoning.  
XX Small round structured virus.  
XX WO200079280-A1.  
XX 28-DEC-2000.  
XX 22-JUN-2000; 2000WO-JP004095.  
XX 22-JUN-1999; 99JP-00175928.  
XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.  
PA (DENK-) DENKA SEIKEN KK.  
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;  
XX WPI; 2001-080848/09.  
DR N-PSDB; AAF29147.  
XX Kit for the detection and typing of small round-structured virus (SRSV)  
PT strains for investigation of food poisoning outbreaks, contains  
PT antibodies.  
XX Claim 1; Page 54-57; 84pp; Japanese.  
XX This invention relates to a kit for the detection and typing of small  
CC round structured virus (SRSV) strains. The kit contains antibodies  
CC directed against peptides represented in sequences AAB49700 - AAB49710,  
CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -  
CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is  
CC used for detecting and typing strains of SRSV in order to prevent the  
CC spread of infection and to examine the epidemiology of outbreaks  
XX Sequence 540 AA;  
Query Match 57.0%; Score 1650.5; DB 4; Length 540;  
Best Local Similarity 56.5%; Pred. No. 3.7e-143;  
Matches 330; Conservative 65; Mismatches 102; Indels 87; Gaps 8;  
QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVQAP 60  
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVQAP 60  
QY 61 AGEFTVSPRNSGGEILLDELGPDLPNLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120  
DB 61 AGEFTVSPRNSGGEILLDELGPDLPNLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120  
QY 121 FAAIPPGPYENLSPSQITMCPHVIIDVROLEPFLPMPDIWNFFHYNQGNPKRLVA 180  
DB 121 FAAIPPGPYENLSPSQITMCPHVIIDVROLEPFLPMPDIWNFFHYNQGNPKRLVA 180  
QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNS 240  
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFEFTFLVPPTVESKTKQFALPILKISEMTNS 240  
QY 241 RPPVPDVMYATARNENQVVPQNGRVLTDGELLGTTPLLAIVNICFKGEVIK-NGDVRS 299



PS Claim 1; Page 62-64; 84pp; Japanese.

XX This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

XX Sequence 550 AA;

Query Match 55.8%; Score 1613.5; DB 4; Length 550;  
Best Local Similarity 55.3%; Pred. No. 1e-139; Mismatches 99; Gaps 12;  
Matches 329; Conservative 77;

QY 1 MKWASNDAAAPSDGAGLVPEINNEVMPLEPVAGASLATPVVGGQNIIDPWIRNPFVQAP 60  
DB 1 MKWASNDAAAPSDGAGLVPEINNEVMPLEPVAGASLATPVVGGQNIIDPWIRNPFVQAP 60

QY 61 AGEFTVSPRNSPGEILLDELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120  
DB 61 AGEFTVSPRNSPGEILLDELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120

QY 121 FPAIPGPPYENLSQITMCHVVIDVROLEPFLPMPDIWNNFHHYNGNDPKLRLVA 180  
DB 121 FPAIPGPPYENLSQITMCHVVIDVROLEPFLPMPDIWNNFHHYNGNDPKLRLVA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDFEFTLVPPVTSKTKQFALPILKISEMTNS 240  
DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDFEFTLVPPVTSKTKQFALPILKISEMTNS 240

QY 241 RFPPVDMYMTARNENQVVPQNGRVTLDGELLGTTPLLAVNICKFKGEVIARNG 295  
DB 241 RFPPVDMYMTARNENQVVPQNGRVTLDGELLGTTPLLAVNICKFKGEVIARNG 295

QY 296 -----DVRSYRMEITNDGTPIPTDPTDPTGPIGSPDQGIILFGVASORNEQNP 347  
DB 296 -----DVRSYRMEITNDGTPIPTDPTDPTGPIGSPDQGIILFGVASORNEQNP 347

QY 300 TDSQARAHNHLVQKNDLQYDPTDDIPAVLGAIDPKGTGVFGVASQRDVSGQOEGH 359  
DB 300 TDSQARAHNHLVQKNDLQYDPTDDIPAVLGAIDPKGTGVFGVASQRDVSGQOEGH 359

QY 348 -ATRAHEALINTGGDHLCPQISSEIYLTSPNLCRTNPQPLPQSLGRTILIRSDNGHC 406  
DB 348 -ATRAHEALINTGGDHLCPQISSEIYLTSPNLCRTNPQPLPQSLGRTILIRSDNGHC 406

QY 360 YATRAHEALINTGGDHLCPQISSEIYLTSPNLCRTNPQPLPQSLGRTILIRSDNGHC 387  
DB 360 YATRAHEALINTGGDHLCPQISSEIYLTSPNLCRTNPQPLPQSLGRTILIRSDNGHC 387

QY 407 HDVMTGSPITPTWQPMRRCRSGNSCCSGHRYPPVPMNRTM 450  
DB 407 HDVMTGSPITPTWQPMRRCRSGNSCCSGHRYPPVPMNRTM 450

QY 388 -----GSDDFTNQPIRFTVPGNGDN---NWRQWELPDYSGRLTLNNLAPAVSPFPCR 440  
DB 388 -----GSDDFTNQPIRFTVPGNGDN---NWRQWELPDYSGRLTLNNLAPAVSPFPCR 440

QY 451 -----IVLS---HKSFGSTSTRKLPQLNLRW-----PLIRFINDPTGRV 486  
DB 451 -----IVLS---HKSFGSTSTRKLPQLNLRW-----PLIRFINDPTGRV 486

QY 441 ILFFRSIVPSAGYSGYIDCL--IPQ---EWWQHFYQEAAPSQSAVALVRYVNPDTGRN 495  
DB 441 ILFFRSIVPSAGYSGYIDCL--IPQ---EWWQHFYQEAAPSQSAVALVRYVNPDTGRN 495

QY 487 LFEARLHKOGFTVAHTGDNPIVMPNGYFRFEAMVNOFYSLAPVCTGKGRRRVQ 541  
DB 487 LFEARLHKOGFTVAHTGDNPIVMPNGYFRFEAMVNOFYSLAPVCTGKGRRRVQ 541

QY 496 IFEAKLHREGFTLVANGNPIVPPNGYFRFEAMVNOFYSLAPVCTGKGRRRVQ 550  
DB 496 IFEAKLHREGFTLVANGNPIVPPNGYFRFEAMVNOFYSLAPVCTGKGRRRVQ 550

RESULT 7

AAB49704

ID AAB49704 standard; protein; 539 AA.

XX AAB49704;

XX 04-APR-2001 (first entry)

XX Small round structured virus protein SEQ ID 5.

DE Small round structured virus; SRSV; food poisoning.

XX Small round structured virus.

OS Small round structured virus.

XX WO200079280-A1.

FN

XX

PD 28-DEC-2000.

XX 22-JUN-2000; 2000WO-JP004095.

XX 22-JUN-1999; 99JP-00175928.

PR (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

PA (DENK-) DENKA SEIKEN KK.

XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

PI WPI: 2001-080848/09.

XX N-PSDB; AAF29145.

DR Kit for the detection and typing of small round-structured virus (SRSV) strains for investigation of food poisoning outbreaks, contains antibodies.

PT Claim 1; Page 50-52; 84pp; Japanese.

PS This invention relates to a kit for the detection and typing of small round structured virus (SRSV) strains. The kit contains antibodies directed against peptides represented in sequences AAB49700 - AAB49710, which are each SRSV strain specific. Polynucleotide sequences AAF20141 - AAF20151 represent cDNA encoding the strain specific proteins. The kit is used for detecting and typing strains of SRSV in order to prevent the spread of infection and to examine the epidemiology of outbreaks

XX Sequence 539 AA;

QY 1 MKWASNDAAAPSDGAGLVPEINNEVMPLEPVAGASLATPVVGGQNIIDPWIRNPFVQAP 60  
DB 1 MKWASNDAAAPSDGAGLVPEINNEVMPLEPVAGASLATPVVGGQNIIDPWIRNPFVQAP 60

QY 61 AGEFTVSPRNSPGEILLDELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120  
DB 61 AGEFTVSPRNSPGEILLDELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120

QY 121 FPAIPGPPYENLSQITMCHVVIDVROLEPFLPMPDIWNNFHHYNGNDPKLRLVA 180  
DB 121 FPAIPGPPYENLSQITMCHVVIDVROLEPFLPMPDIWNNFHHYNGNDPKLRLVA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDFEFTLVPPVTSKTKQFALPILKISEMTNS 240  
DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDFEFTLVPPVTSKTKQFALPILKISEMTNS 240

QY 241 RFPPVDMYMTARNENQVVPQNGRVTLDGELLGTTPLLAVNICKFKGEVIARNGDVRSY 300  
DB 241 RFPPVDMYMTARNENQVVPQNGRVTLDGELLGTTPLLAVNICKFKGEVIARNGDVRSY 300

QY 301 RMDMEITNDGTPIPTDPTDPTGPIGSPDQGIILFGVASORNEQNPATRAHEALINTGG 360  
DB 301 RMDMEITNDGTPIPTDPTDPTGPIGSPDQGIILFGVASORNEQNPATRAHEALINTGG 360

QY 298 DYTNNLASQWMSNYDPTTEEIPAPLGTDPFVGIQGLTQTR--EDGSTRAHKATVSTGS 355  
DB 298 DYTNNLASQWMSNYDPTTEEIPAPLGTDPFVGIQGLTQTR--EDGSTRAHKATVSTGS 355

QY 361 DHLCPQISSEIYLTSPNLCRTNPQPLPQSLGRTILIRSDNGHCHDMVGSPTPTTPW 420  
DB 361 DHLCPQISSEIYLTSPNLCRTNPQPLPQSLGRTILIRSDNGHCHDMVGSPTPTTPW 420

QY 356 VHTFPLKLSVQ--YTTDTN-----NDFTQNTKFTPVGVIQDGNHNE-----P 399  
DB 356 VHTFPLKLSVQ--YTTDTN-----NDFTQNTKFTPVGVIQDGNHNE-----P 399

QY 421 QOWRRCRSGNSCCSGHRYPPVPMNRTMIVLSHKSFGSTSTRKLPQLNL-----R 472  
DB 421 QOWRRCRSGNSCCSGHRYPPVPMNRTMIVLSHKSFGSTSTRKLPQLNL-----R 472

QY 400 QOWVLPNYSGRTGHNVLAPAVAPTFPGEQLLFFESTWPGCS---GYPMNMLDCLLPDE 455  
DB 400 QOWVLPNYSGRTGHNVLAPAVAPTFPGEQLLFFESTWPGCS---GYPMNMLDCLLPDE 455

QY 473 W-----PLIRFINDPTGRVLFEARLHKOGFTVAHTGDNPIVMPNGYFRFEAMVNOFYSLAPVCTGKGRRRVQ 517  
DB 473 W-----PLIRFINDPTGRVLFEARLHKOGFTVAHTGDNPIVMPNGYFRFEAMVNOFYSLAPVCTGKGRRRVQ 517

QY 456 WYQHFCEQAAPASQVALLRFVNPDTGRVLFCKLHKSQYVTVANTGPHDLVIPNGYFR 515  
DB 456 WYQHFCEQAAPASQVALLRFVNPDTGRVLFCKLHKSQYVTVANTGPHDLVIPNGYFR 515

QY 518 FEAMVNOFYSLAPVCTGKGRRR 539  
DB 518 FEAMVNOFYSLAPVCTGKGRRR 539

QY 516 FDSWYNQYFTLAPMGNGAGRRR 537  
DB 516 FDSWYNQYFTLAPMGNGAGRRR 537

Db 301 NL-LQLTPNGASYDPTDEVPAFLGTQDFSGMLYGVLTQDNVNVSTGEAKNAKGIYSTT 359

Qy 360 GDHLCPOISSSEIYLTSPNILACTNPQPLPQSLRGTLIRSDNGHCHDMV----- 410

Db 360 SGKFTPKIGSIGLHSITEHV-----HPNQSRFTPPGVAVDENTPFOQWVLPVHAGSLA 413

Qy 411 ---GTSPT-TTPTWPOQWRRCSRSGNSCCSGHRYPPVVMNRVTWVLSHSGSFSTSKL 466

Db 414 LNTNLAPAVAPTFPGEOQLLFFRSRVPCVG-----LQQDAFIDCL--L 455

Qy 467 POLNLRW-----PLIRFNPDTGRVLFEARLHKQGFITVAHTGDNPIVMP 511

Db 456 PQ---EWNHIFYQEAAPSQADVALLIRYVNPDTGRTLFEAKLHRSFGFITVSHGTGALPVPV 512

Qy 512 PNGYFRFEAWNQFYSLAPVGTGKGRRYQ 541

Db 513 PNGHFRFDSWVNOFYSLAPMGTGNGRRRIQ 542

RESULT 9

AAB49700 standard; protein; 545 AA.

XX AAB49700;

AC AAB49700;

XX 04-APR-2001 (first entry)

DT Small round structured virus protein SEQ ID 1.

DE Small round structured virus; SRSV; food poisoning.

KW Small round structured virus.

OS Small round structured virus.

XX WO200079280-A1.

PN 28-DEC-2000.

PD 22-JUN-2000; 2000WO-JP004095.

PF 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

PR (DENK-) DENKA SEIKEN KK.

PA Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;

PI WPI; 2001-080848/09.

DR N-PSDB; AAF29149.

XX Kit for the detection and typing of small round-structured virus (SRSV)

PT strains for investigation of food poisoning outbreaks, contains

PT antibodies.

XX Claim 1; Page 59-61; 84pp; Japanese.

PS This invention relates to a kit for the detection and typing of small

XX round structured virus (SRSV) strains. The kit contains antibodies

CC directed against peptides represented in sequences AAB49700 - AAB49710,

CC which are each SRSV strain specific. Polynucleotide sequences AAF20141 -

CC AAF20151 represent cDNA encoding the strain specific proteins. The kit is

CC used for detecting and typing strains of SRSV in order to prevent the

CC spread of infection and to examine the epidemiology of outbreaks

XX Sequence 542 AA;

Qy Query Match 53.2%; Score 1538.5; DB 4; Length 542;

Db Best Local Similarity 54.4%; Pred. No. 8.4e-133; Indels 57; Gaps 9;

Qy Matches 310; Conservative 72; Mismatches 131;

Db 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQGNIIIDPWRNFVQAP 60

Qy 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQGNIIIDPWRNFVQAP 60

Qy 61 AGEFTVSPNSPGEILLDLELGPDLNPLYLAHRYNCHAGMEVQIVLAGNAFTAGKII 120

Db 61 NGEFTVSPNSPGEILLDLELGPDLNPLYLAHRYNCHAGMEVQIVLAGNAFTAGKII 120

Qy 121 FRAIPGFPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNFFHYNQGNPKLRJVA 180

Db 121 FRAIPGFPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNFFHYNQGNPKLRJVA 180

Qy 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFETFLVPPVESKTKQFALPKIKSEMTNS 240

Db 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFETFLVPPVESKTKQFALPKIKSEMTNS 240

Qy 241 RFPVVDVMTARENQVQPNQGRVTLDELGLLTPPLAVNICPKFGEVIAKNGDVRSY 300

Db 241 RFPPIEQLYATPNETNVQCNQGRCTLDGELGTTQLLSAVCFLOGRTVADNGNDWQ 300

Qy 301 RMDMEITNDGTPTIDPTEDTFPGIPSGDFQGLIFGVASQRNKN-EQNPAATRAHEALINTG 359

QY 57 VQAPAGEFTVSPRNSPGEILLDLELGPDLNPNYLAHLARMYNGHAGMEVOIVLAGNAFTA 116  
 DB 61 VQAPAGEFTISPNTTSGDILFDLQGLPHLNPFLSHLAQMYNGWGVNKKVLAGNAFTA 120  
 QY 117 GKIIIPAAIPGPPYENLSPQITMCPHVLIIDVQLEPFLPMPDIWNNEFFHYNOGNDPKL 176  
 DB 121 GKIIISCIPGFAAQAATMFAHVIADVRVLEPVEDVRNLVLFH--NNDNAPTM 179  
 QY 177 RLVMALYTLPLRA--NNSGDDVFTVSCRVLTKPSDFEFTLVPPPTVESKTKOFALPILKI 234  
 DB 180 RLVMALYTLPLRASGSSGTDPPFIAGRVLTCPSPDFSLFLVPPNVEQTKPSPVENLPL 239  
 QY 235 SEMTNSRFPVVDVMTARNENQVQONGRVTLDELGLGTTPLLAVINCKFKGEVIAQN 294  
 DB 240 NTLNSRVPFLIKSMVMVRDHGMQVQNGRVTLGQLOQTTPTSASQLCKIRGSVFHAN 299  
 QY 295 GIVRSYRMDMEINTDGTPTDPTEDTPGIPSGDFQILFGVASQRNKNQONPATRAHEA 354  
 DB 300 GG-NGY----NLTELDGSPYHAF--SPAFIGPFDLGECDWHM-----EASFTTQ---- 342  
 QY 355 IINTGGDLCPQISSEIYLTSPNLRCTNPOPLQSGLRGTILI--RSDNGHCHDMVG- 411  
 DB 343 -FNTG--DIVIKQINVKQESAFAPHL-----GTIQADGLSDVSNTNMIK 384  
 QY 412 ----TSPTTTPWQOQWRRRCRSGNCCSGHRYPPVVMNRVTWIVLSHKSGFSTSTRKLPQ 468  
 DB 385 LGWVSPV-----SDGHRGDVD-----PWVPIPRYGSTLTTEAAQLAPP 420  
 QY 469 L-----NLWLP-----LIRFINPD 482  
 DB 421 IYPPGGEIAIVFMSDFPIAHGTNGLSVPTCIPQEFVTHFVNEQAFTRGEAALLHYLDDP 480  
 QY 483 TGRVLFEARLHKQGFITVA--HTGDNPIVMPPNGYFRFEAWNOFYSLAPVGT 533  
 DB 481 THRNLFGEFLYEGEFTVCPNNSGTCGPQLPENGVFVFSWSRFLQALPVG 533

RESULT 10  
 AAR50972  
 ID AAR50972 standard; protein; 530 AA.  
 XX AC AAR50972;  
 XX DT 16-OCT-2003 (revised)  
 DT 25-MAR-2003 (revised)  
 DT 05-OCT-1994 (first entry)  
 XX DE Norwalk virus strain 8FIIa protein (encoded by ORF2).  
 XX KW Norwalk virus; pathogen; acute gastroenteritis; food poisoning;  
 KW seafood contamination; diagnostic assay; calcivirus; small round virus.  
 XX OS Norwalk virus; (strain 8FIIa).  
 XX PN W09405700-A2.  
 XX PD 17-MAR-1994.  
 XX PF 07-SEP-1993; 93WO-US008447.  
 XX PR 07-SEP-1992; 92US-00941365.  
 XX PA (BAYU ) BAYLOR COLLEGE MEDICINE.  
 XX PI Matson DO, Estes MK, Jiang X, Graham DY;  
 XX WPI; 1994-101125/12.  
 DR N-PSDB; AAQ56826.  
 XX DNA from Norwalk and related viruses - used for preparing prods. for use  
 PT in diagnostic assays, detection and vaccines for Norwalk and related  
 PT viruses.

XX Claim 14; Page 68-70; 156pp; English.  
 XX The Norwalk virus was isolated from stool samples from adult volunteers  
 CC infected with safety tested Norwalk virus strain 8FIIa. The coding  
 CC sequence is useful for the design of probes for use in diagnostic assays  
 CC for the Norwalk and related viruses. (Updated on 25-MAR-2003 to correct  
 CC PN field.) (Updated on 16-OCT-2003 to standardise OS field)  
 XX SQ Sequence 530 AA;  
 Query Match 34.8%; Score 1008.5; DB 2; Length 530;  
 Best Local Similarity 40.9%; Pred. No. 7.4e-84;  
 Matches 235; Conservative 72; Mismatches 180; Indels 87; Gaps 14;  
 QY 1 MKVASNDAAAPSSDGAAG---LVPEIN-NEVMPLBEPVAGASLATPVVGOQNIIDPWRNPF 56  
 DB 1 MMWASKDATSSVDGASGAGQLVPEVNASDPLAMDVPAGSSTAVATAGVQNPIDPWIINPF 60  
 QY 57 VQAPAGEFTVSPRNSPGEILLDLELGPDLNPNYLAHLARMYNGHAGMEVOIVLAGNAFTA 116  
 DB 61 VQAPAGEFTISPNTTSGDILFDLQGLPHLNPFLSHLAQMYNGWGVNKKVLAGNAFTA 120  
 QY 117 GKIIIPAAIPGPPYENLSPQITMCPHVLIIDVQLEPFLPMPDIWNNEFFHYNOGNDPKL 176  
 DB 121 GKIIISCIPGFAAQAATMFAHVIADVRVLEPVEDVRNLVLFH--NNDNAPTM 180  
 QY 177 RLVMALYTLPLRASGSSGTDPPFIAGRVLTCPSPDFSLFLVPPNVEQTKPSPVENLPL 236  
 DB 180 RLVMALYTLPLRASGSSGTDPPFIAGRVLTCPSPDFSLFLVPPNVEQTKPSPVENLPL 240  
 QY 237 MTNSRFPVVDVMTARNENQVQONGRVTLDELGLGTTPLLAVINCKFKGEVIAKNGD 296  
 DB 241 LNSRAPLPISSMGISPDNVQSVQFQNGRCTLGRLVGTTPVSLSHVAKIRG---TSNGT 297  
 QY 297 VRSYRMDMEINTDGTPTDPTEDTPGIPSGDFQILFGVASQRNKNQONPAT 349  
 DB 298 V-----INLTEDGTTPHFFEG--PAPIGFPLDGLGCDWHMNTQFGHSSQTOYD----- 344  
 QY 350 RAHEAIINTGGDLCPQISSEIYLTSPNLRCTNPOPLQSGLRGTILIIRSDNGHCHDM 409  
 DB 345 -----VDTPPTFPVHLGS-----IQANGTSGNY 369  
 QY 410 VG-----TSPTTTPWQ--QMRRCRSGNCCSGHRYP--VPVVMNRVTWIVLSHKSGFS 460  
 DB 370 VGVLSWISGPPSPHSGSQVDLWKIPNYGSSITEATHLAPSVYPPGGEVLVFFMSPKMPGP 429  
 QY 461 TSTRK--LPQLNL-----RWPLIRFINPDGTGRVLFEARLHKQGFITVAHTG-- 504  
 DB 430 AYNLPCLLPQEIYISHLASEQAPTVGEAALLHYVDPDTCRNLGFEKAYPDGFLTCVPNGAS 489  
 QY 505 DNPIVMPPNGYFRFEAWNOFYSLAPVGTGKRR 538  
 DB 490 SGFQQLPINGVFVFSWSRFLQALPVGTTASSAR 523

RESULT 11  
 ADC72176  
 ID ADC72176 standard; protein; 530 AA.  
 XX AC ADC72176;  
 XX DT 18-DEC-2003 (first entry)  
 XX DE Norwalk virus protein 2 amino acid sequence.  
 XX KW immune response; non-Norwalk virus agent; immunogen; Norwalk virus;  
 KW viral pathogen; acute gastroenteritis; virucidal; antiviral vaccine.  
 XX OS Norwalk virus.  
 XX PN US6572862-B1.  
 XX

```
PD 03-JUN-2003.
XX
PF 07-JUN-1995; 95US-00486049.
XX
XX 08-NOV-1989; 89US-00433492.
PR 27-APR-1990; 90US-00515993.
PR 27-AUG-1990; 90US-00573509.
PR 06-MAY-1991; 91US-00696454.
XX
PA (BAYU ) BAYLOR COLLEGE MEDICINE.
XX
PI Bates MK, Jiang X, Graham DY;
XX
XX WPI; 2003-776005/73.
XX
DR N-PSDB; ADC72174.
XX
XX Inducing an immune response against non-Norwalk virus agents, comprises
XX administering an immunogen recombinantly expressed from a cDNA from
XX Norwalk virus.
XX
XX Example 4; SEQ ID NO 3; 45pp; English.
XX
XX This invention relates to a novel method of inducing an immune response
XX in an individual against Norwalk virus and non-Norwalk virus agents, by
XX orally or parenterally administering an immunogen recombinantly expressed
XX or synthesised from a cDNA of Norwalk virus given in the specification.
XX CC Norwalk virus is one of the most important viral pathogens, causing acute
XX gastroenteritis. The invention may be used for the development of
XX compounds with virucidal activity or an antiviral vaccine. The present
XX sequence is the amino acid sequence of a protein encoded by the Norwalk
XX virus genome of the invention.
XX
XX Sequence 530 AA;
XX
XX Query Match 34.8%; Score 1008.5; DB 7; Length 530;
XX Best Local Similarity 40.9%; Pred. No. 7.4e-84;
XX Matches 235; Conservative 72; Mismatches 180; Indels 87; Gaps 14;
XX
XX 1 MKMASNDAAAPSSDGAAG---LVPEIN-NEVMPELPVAGASLAPVVGQNIIDPWIRNNF 56
XX 1 MMASKADATSSVDGASGAGQLVPEVNASDPLAMDPLMDPVAGSSTAVATAGVNPIDPWIRNNF 60
XX
XX 57 VOAPAGEFTVSPRNSPGEILLDLGLPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTA 116
XX 61 VOAPQGEFTISPNNTPGGVLFDSLGLPHLNPFLHLSQMYNGWGNRVRIMLAGNAFTA 120
XX
XX 117 GKIIFAAIPPGPFYENLSPSQITMCPHVIDVRLQLEPFLPMPDIWNNFFHYNQNDPKL 176
XX 121 GKIIIVSCIPPGFGSHNLTTAQTLPFHVIADVRLDPIEVLDPVRLVFNHNDNRNQTM 180
XX
XX 177 RLVMALYTPLRANNSGDDVFTVSCRVLTKPSDREFTFLVPPTVESKTKQFALPKIKISE 236
XX 181 RLVCMLYTPLRGTGGTGDSFVAGRWMTCPSPDFNPLFLVPTVEQKTRPTLPLNPLSS 240
XX
XX 237 MTNSRFPVPVDMYTARNENQVQPQNGRVTLDGELLGTPPLLANVICKFGEVIKNGD 296
XX 241 LNSRAPLFISSMGISPDNVQSQVQFQNGRCTLGRLVGTTPVSLSHVAKIRG---TSNGT 297
XX
XX 297 VRSYRMDMEITNDGTPIPTDPTDTPGIGSPDFQG-----ILFGVASQRNKEONPAT 349
XX 298 V-----INLTLDGTPFPFPEG-PAPIGPDLGGCDWHNNITQFGHSQTOYD----- 344
XX
XX 350 RAHEALINTGDHLCPOISSEIYLTSPNILRCTNPQLPQSLGRTILIRSDNGHCHDM 409
XX 345 -----VDITPDTFVPHLGS-----IQANGIGSGNY 369
XX
XX 410 VG-----TSPTTPTWFO--QWRRCRSGSNCCSGHRYP--VPVMNRVTWIVLSHKSQFS 460
XX 370 VGVLSWISPPSPHSPSGSQVLDWKIPNYGSSITEATHLAPSYYPGFGFVLVFFWMSKMPBG 429
XX
XX 461 TSTRK--LPQLN-----RWPLIRFINPDGTGRVLFPEARLHKQGFITVAHTG-- 504
XX 430 AYNLPCLLPQEIYSHLASEQAPTVGEAALLHYVDPTGRNLGFEKAYPDGFLTCVDPNGAS 489
```



```

Db      241  LNSRAPLPISGMGISSPDNVQSFQNGRCTLGRLVGTTPVSLSHVAKIRG---TSNGT 297
QY      297  VRSYRMDMEITNDGTPIDPTEDTPGPIGSPDFQG-----ILFGVASQRNKNQNPAT 349
Db      298  V-----INLTDLGTPFPFPEG-PAPIGFDPDLGGCDWHINMTQFGHSSQTQYD----- 344
QY      350  RAHEAIINTGGHLCPOISS-----SEIYUTSPNLRCTNPQPLPQSGGLRTILIRSD 402
Db      345  -----VDTPDTFVPHLGSIQANGIGSGNYI---GVLSWVSP----- 378
QY      403  NGCHDMVGTSTPTTPWQ--QWRRCRSGSNCCSSGHRYP--VPVVMNRVTWVLSHKSQ 458
Db      379  -----PSHSGSQVDLWKIPNYGSSITEATHLAPSVPYPGFGEVLVFFSKIPG 427
QY      459  FSTSTRK--LPQNLN-----RWPLRFINPDTGRVLFEARLHKQGFITVAHTG 504
Db      428  PGAYSPLCLLPQEIYISHLASEQAPTVEAALLHYVDPDTGRTLGEPKAYPDGFLTCVPNG 487
QY      505  --DNPIVMPNGYFRFEAWNQFYSLAPVGTGKRR 538
Db      488  ASSGQQLPINGVFVFSWVSRYQLKPVGTASSAR 523

```

## RESULT 13

ID AAR57091 standard; protein; 530 AA.

AC AAR57091;

DT 27-AUG-2003 (revised)

DT 25-MAR-2003 (revised)

DT 05-OCT-1994 (first entry)

XX Small round virus SRSV/KY/89 capsid protein.

XX pathogen; acute gastroenteritis; food poisoning; seafood contamination;

KW diagnostic assay; human calicivirus; small round virus; SRSV; KY89;

KW Norwalk virus; capsid protein.

XX Small round structured virus.

XX WO9405700-A2.

XX 17-MAR-1994.

XX 07-SEP-1993; 93WO-US008447.

XX 07-SEP-1992; 92US-00941365.

XX (BAYU ) BAYLOR COLLEGE MEDICINE.

XX Matson DO, Estes MK, Jiang X, Graham DY;

XX WPI; 1994-101125/12.

DR N-PSDB; AAQ56832.

PT DNA from Norwalk and related viruses - used for preparing prods. for use  
PT in diagnostic assays, detection and vaccines for Norwalk and related  
PT viruses.

XX Example 7; Fig 13a; 156pp; English.

XX The known sequence for Norwalk virus was used to obtain the sequence of  
CC other Norwalk-related viruses such as SRSV/KY/89, an agent from a stool  
CC from an outbreak of gastroenteritis in Japan in 1989. The 2516 nucleotide  
CC cDNA sequence includes part of the polymerase region and the capsid  
CC region of the genome; the deduced amino acid sequences are AAR57092 and  
CC AAR57091, respectively. Expression of fragments and derive. of Norwalk-  
CC related viruses permits development of diagnostic assays to detect  
CC antibodies, antigens, viral genetic material or antivirals. (Updated on  
CC 25-MAR-2003 to correct FN field.) (Updated on 27-AUG-2003 to correct OS  
CC field.)

XX SQ Sequence 530 AA;

Query Match 34.0%; Score 984.5; DB 2; Length 530;

Best Local Similarity 39.3%; Pred. No. 1.2e-81;

Matches 230; Conservative 78; Mismatches 168; Indels 109; Gaps 16;

QY 1 MKGASNDAAPSDGAAG---LVPEIN-NEVMPLEPVAGASLATPVVQGNIIIDPWIRNPF 56

Db 1 MMASKDATSSVDGASASVQLVPEVNASDPLAMDPAVAGSTAVATAGVQNPIDPWIINF 60

QY 57 VOAPAGEFTVSPRNSPGEILDLGLGPDNLNVLALHARMYNGHAGMEVQIVLAGNAFTA 116

Db 61 VOAPGGEFTISNNTPGDVLFDLSUGPHLNPFLHLSQMYNGWGMVRIMLAGNAFTA 120

QY 117 GKIIIFAAIPPGFPYENLSPSQITMCPHVIIDVQLEPFLPMPDIWNNEFFHYNQGNPKL 176

Db 121 GKIIIVSCIPPGFGSQLTIAQATLFPHVIAVDTLDPIEVLEDVRNVLPHNDRNQTM 180

QY 177 RLVMALYTPLRANSGDDVFTVSCRVLTKPSDPDEFTFLVPTVESKTKQPALPIKISE 236

Db 181 RLVCMLYTPLSTGGTGDSFVVAGVMTCPSPDFNPLFLVPTVEQKTRPFTLPNPLSS 240

QY 237 MTNSRFPPVVDVMTARNENQVQPQNGRVTLDGELGTTPLLAVNICFKGEVIKNGD 296

Db 241 LNSRAPLPISGMGISSPDNVQSFQNGRCTLGRLVGTTPVSLSHVAKIRG---TSNGT 297

QY 297 VRSYRMDMEITNDGTPIDPTEDTPGPIGSPDFQG-----ILFGVASQRNKNQNPAT 349

Db 298 V-----INLTDLGTPFPFPEG-PAPIGFDPDLGGCDWHINMTQFGHSSQTQYD----- 344

QY 350 RAHEAIINTGGHLCPOISS-----SEIYUTSPNLRCTNPQPLPQSGGLRTILIRSD 402

Db 345 -----VDTPDTSPVHLGSIQANGIGSGNYI---GVLSWVSP----- 378

QY 403 NGCHDMVGTSTPTTPWQ--QWRRCRSGSNCCSSGHRYP----- 440

Db 379 -----PSHSGSQVDLWKIPNYGSSITEATHLAPSVPYPGFGEVLVFFSKIPG 427

QY 441 -----VPVVMNRVTWVLSHKSQFSTSTRKLPQLNLRWPLRFINPDTGRVLFEARLHKQ 495

Db 428 PGDLSPLCLLPQG---YISH-----LASEQPTVG-EGPFLHYVDPDTRNLGEFKAYPD 478

QY 496 GFITVAHTG--DNPIVMPNGYFRFEAWNQFYSLAPVGTGKRR 538

Db 479 GFLTCVPNGASSGPOQLPINGVFVFSWVSRYQLKPVGTASTAR 523

## RESULT 14

AAB49703

ID AAB49703 standard; protein; 544 AA.

AC AAB49703;

XX 04-APR-2001 (first entry)

DT Small round structured virus protein SEQ ID 4.

DE Small round structured virus; SRSV; food poisoning.

KW Small round structured virus.

XX Small round structured virus.

XX WO200079280-A1.

XX 28-DEC-2000.

PF 22-JUN-2000; 2000WO-JP004095.

XX 22-JUN-1999; 99JP-00175928.

XX (NINA-) JAPAN NAT INST INFECTIOUS DISEASES.

PA (DENK-) DENKA SEIKEN KK.

XX

```
PI Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
XX WPI: 2001-080848/09.
DR N-PSDB; AAF29144.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
XX Claim 1; Page 47-49; 84pp; Japanese.
PS
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 544 AA;
SQ
Query Match 33.5%; Score 970; DB 4; Length 544;
Best Local Similarity 38.0%; Pred. No. 2.8e-80;
Matches 224; Conservative 91; Mismatches 176; Indels 98; Gaps 17;
QY 1 MKNASNDAAPSDGAAG---LVPEINN-EVMPLEPVAGASLATPVVQGNIIIDPWIRNMF 56
Db 1 MWMASKDAPSDGATGAGQVPEVNTADPIPIIDPVAGSSTALATAGQVNLIDPWIRNMF 60
QY 57 VQAPAGEFTVSPRNSGCEIILDLGLDLPNLYLAHARMYNGHAGMEVQIVLAGNAFTA 116
QY 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGMVRVVLVLAGNAFTA 120
QY 117 GKIIIFAAIPGFPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNFFHYNQNDPKL 176
Db 121 GKVIICCVPGFQSRISIAQATLPHVIADVRITLDPVEVPLEDVRNVLVH-NNDTQPTM 179
QY 177 RLVAWLYTPILAN--NSGDDVFTVSCRVLTKPSDPDEFTELVPPTVESKTKQFALPILKI 234
Db 180 RLCLMLYTPILRTGGASGSDTSFVAGRVLTCPGDFNLFVLVPTVEQKTRPTVFNIP 239
QY 235 SEMTNSRFPVVDVMTARNENQVQNGRVTLTGELGTLTPILLAVNICKFGEVIAKN 294
Db 240 KYLSNRLIPNFIEMSLSPDQTVQVQNGRCTIDGQPLGTTTPVSQVQLCKFRGR---- 294
QY 295 GDVRSYRMDMEITNTDGTPTDPTEDTPGIPGSPDFQGLIFGVASQRNKN---EQNP---- 347
Db 295 --ITSQVRVNLTELDSGPF--MAFAAPAPAGFPDLGSCDWHIEMSKIENSTQNNPIVTN 351
QY 348 -----ATRAHEALINTGDLHCPQISSEIYLTSPNILRCTNPNQPLPQSGLR 394
Db 352 SVKPNSSQFVPHLSSITLDENVSSGDYI-----GTIQWTSF----- 388
QY 395 GTILIRSDNGHCHDMVGTSTPTTPWQWRRCRSGNCCSSGHRYPV--PVMNRVTWIV 452
Db 389 -----PSUGGAN-----TNFKIPDYGSSLAELASQAPVYPGFNEVIVYF 431
QY 453 LSHKSGFSTSTRK-----LQO-----LNLWPF-----LIRPNPTGRLPEARLHK 494
Db 432 MASIFGPNQSGSNLVPCLLPQETITHFISEQAPITQGEAALLHYVDPDTNRNLGFEKLYP 491
QY 495 QGFITVA--HTGDNPIVPPNPGYFRFEAWNOFYSLAPVGT--GKGRREV 540
Db 492 GGYLTCVPSNSSSTGQQQLDGVFVFASWVSRYQLKPVGTAGPARGRL 540
RESULT 15
ID AAB49702
AC AAB49702 standard; protein; 546 AA.
XX AAB49702;
XX
XX 04-APR-2001 (first entry)
```

```
XX Small round structured virus protein SEQ ID 3.
DE
XX Small round structured virus; SRSV; food poisoning.
KW
XX Small round structured virus.
OS
XX W0200079280-A1.
PN
XX 28-DEC-2000.
PD
XX 22-JUN-2000; 2000WO-JP004095.
PF
XX 22-JUN-1999; 99JP-00175928.
PR
XX (NTNA-) JAPAN NAT INST INFECTIOUS DISEASES.
PA (DENK-) DENKA SEIKEN KK.
XX
XX Takeda N, Natori K, Miyamura T, Kamata K, Sato T, Sato S;
PI
XX WPI: 2001-080848/09.
DR N-PSDB; AAF29143.
XX
XX Kit for the detection and typing of small round-structured virus (SRSV)
PT strains for investigation of food poisoning outbreaks, contains
PT antibodies.
XX
XX Claim 1; Page 45-47; 84pp; Japanese.
PS
XX This invention relates to a kit for the detection and typing of small
XX round structured virus (SRSV) strains. The kit contains antibodies
XX directed against peptides represented in sequences AAB49700 - AAB49710,
XX which are each SRSV strain specific. Polynucleotide sequences AAF20141 -
XX AAF20151 represent cDNA encoding the strain specific proteins. The kit is
XX used for detecting and typing strains of SRSV in order to prevent the
XX spread of infection and to examine the epidemiology of outbreaks
XX
XX Sequence 546 AA;
SQ
Query Match 33.4%; Score 967; DB 4; Length 546;
Best Local Similarity 38.5%; Pred. No. 5.3e-80;
Matches 225; Conservative 88; Mismatches 188; Indels 84; Gaps 17;
QY 1 MKNASNDAAPSDGAAG---LVPEINN-EVMPLEPVAGASLATPVVQGNIIIDPWIRNMF 56
Db 1 MWMASKDAPSDGATGAGQVPEVNTADPIPIIDPVAGSSTALATAGQVNLIDPWIRNMF 60
QY 57 VQAPAGEFTVSPRNSGCEIILDLGLDLPNLYLAHARMYNGHAGMEVQIVLAGNAFTA 116
Db 61 VQAPQGEFTISPNNTPGDVLFDLQGLPHLPFLSHLSQMYNGWGMVRVVLVLAGNAFTA 120
QY 117 GKIIIFAAIPGFPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNFFHYNQNDPKL 176
Db 121 GKVIICCVPGFQSRISIAQATLPHVIADVRITLDPVEVPLEDVRNVLVH-NNDTQPTM 179
QY 177 RLVAWLYTPILR--ANNSGDDVFTVSCRVLTKPSDPDEFTELVPPTVESKTKQFALPILKI 234
Db 180 RLCLMLYTPILRTGGSGNSDSFVAGRVLTASDSFSLVPTVEQKTRPTVFNIP 239
QY 235 SEMTNSRFPVVDVMTARNENQVQNGRVTLTGELGTLTPILLAVNICKFGEVIAKN 294
Db 240 QTLNSRFPVVDVMTARNENQVQNGRVTLTGELGTLTPILLAVNICKFGEVIAKN 294
QY 295 GDVRSYRMDMEITNTDGTPTDPTEDTPGIPGSPDFQGLIFGVASQRNKNQNPATRAHEA 354
Db 297 QGART----LNLTEVDGKPF--MAFDSAPVPGFDFGKCDWHMIRISKTPNNTSGDPMRSV 351
QY 355 IINTGDLHCPQISS--SRIYLTSPNILRCTNPNQPLPQSGLRGTLIRSDNGHCHDMVGT 411
Db 352 SVQTNVQGVFPHLGSIQDFEVF-----NHTPDGYIG 382
QY 412 T-----SPTTP--TWQQRRCRSGNCCSSGHRYPVPMNRVTWIVLSHKSGFSTSTR 464
```

Db 383 TIEWISQPTPTDINLWEIPDYSSLSQAANLAP-PVPPPGFGEALVYFVSAPPGNN 441  
Qy 465 K-----LPQ-----LNLRW-----LIRFINPDTGRVLFPEARLHKQGITVA-- 501  
Db 442 RSAPNDVPCLLPQYITHFVSEQPTMGDAALLHYVDPDTNRNLGFEKLYPGGYLTCVEN 501  
Qy 502 HTGDNPIVMPPNGYFRFEAWVQFYSLAPYGTGK-----GRRV 540  
Db 502 GVGAGPQLPLNGVFLFVSWSRFYQLKPVGTASTARSLGVRI 546

Search completed: June 1, 2004, 13:46:03  
Job time : 48.4752 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:37:02 ; Search time 12.8896 Seconds  
(without alignments)  
2166.837 Million cell updates/sec

Title: US-09-926-799-11  
Perfect score: 2894  
Sequence: 1 MKNASNDAPSSDGAAGLVP.....VNFYSLAPVGTGKRRRVQ 541

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/2/iaa/5A COMB.pep.\*  
2: /cgn2\_6/prodata/2/iaa/5B COMB.pep.\*  
3: /cgn2\_6/prodata/2/iaa/6A COMB.pep.\*  
4: /cgn2\_6/prodata/2/iaa/6B COMB.pep.\*  
5: /cgn2\_6/prodata/2/iaa/PCTUS COMB.pep.\*  
6: /cgn2\_6/prodata/2/iaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1008.5	34.8	530	4	US-08-486-049-3
2	267.5	9.2	622	4	US-09-590-020-6
3	265.5	9.2	623	4	US-09-590-020-2
4	265.5	9.2	623	4	US-09-590-020-4
5	264	9.1	668	4	US-09-617-594A-4
6	262.5	9.1	626	4	US-09-590-020-7
7	250.5	8.7	669	4	US-09-617-594A-2
8	114	3.9	630	4	US-09-396-149-5
9	114	3.9	2206	1	US-07-852-260-2
10	114	3.9	2206	1	US-08-461-503-2
11	114	3.9	2206	3	US-08-465-250-2
12	109.5	3.8	1101	3	US-08-986-485-2
13	104.5	3.6	2318	3	US-09-091-219-24
14	104.5	3.6	2318	4	US-09-660-541-24
15	102.5	3.5	596	4	US-09-489-039A-9905
16	101	3.5	1013	3	US-09-415-522-8
17	96	3.3	803	4	US-09-543-681A-4312
18	96	3.3	1082	1	US-08-106-493A-2
19	96	3.3	1082	1	US-08-429-264-2
20	96	3.3	1139	1	US-08-832-883-2
21	96	3.3	1139	2	US-08-832-877-2
22	95	3.3	439	4	US-09-543-681A-5727
23	95	3.3	1050	4	US-09-428-711A-16
24	93.5	3.2	1043	3	US-08-928-361B-30
25	93.5	3.2	1721	3	US-08-700-651-5
26	93.5	3.2	1721	3	US-08-928-361B-6
27	93.5	3.2	1721	4	US-09-588-995A-6

28	93	3.2	1162	3	US-08-803-346-1	Sequence 1, Appl
29	92.5	3.2	1560	4	US-09-264-512B-2	Sequence 2, Appl
30	92	3.2	589	4	US-09-543-681A-6155	Sequence 6155, Ap
31	91	3.1	895	3	US-08-827-962-19	Sequence 19, Appl
32	91	3.1	1162	3	US-08-827-962-15	Sequence 15, Appl
33	90.5	3.1	248	4	US-09-540-236-3073	Sequence 3073, Ap
34	90.5	3.1	1091	3	US-08-986-485-5	Sequence 5, Appl
35	90.5	3.1	1911	4	US-09-854-856-64	Sequence 64, Appl
36	90.5	3.1	1939	4	US-09-854-856-48	Sequence 48, Appl
37	90.5	3.1	1971	4	US-09-854-856-32	Sequence 32, Appl
38	90.5	3.1	1999	4	US-09-854-856-16	Sequence 16, Appl
39	90.5	3.1	2004	4	US-09-854-856-58	Sequence 58, Appl
40	90.5	3.1	2032	4	US-09-854-856-42	Sequence 42, Appl
41	90.5	3.1	2048	4	US-09-854-856-62	Sequence 62, Appl
42	90.5	3.1	2064	4	US-09-854-856-26	Sequence 26, Appl
43	90.5	3.1	2076	4	US-09-854-856-46	Sequence 46, Appl
44	90.5	3.1	2092	4	US-09-854-856-10	Sequence 10, Appl
45	90.5	3.1	2108	4	US-09-854-856-30	Sequence 30, Appl

ALIGNMENTS

RESULT 1  
US-08-486-049-3  
; Sequence 3, Application US/08486049  
; Patent No. 6572862  
; GENERAL INFORMATION:  
; APPLICANT: Bates, Mary K  
; APPLICANT: Jiang, Xi  
; APPLICANT: Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; TITLE OF INVENTION: Characterize No. 6572862walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fulbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE:  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 3:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-486-049-3

Query Match 34.8%; Score 1008.5; DB 4; Length 530;  
Best Local Similarity 40.9%; Pred. No. 4.2e-96;  
Matches 235; Conservative 72; Mismatches 180; Indels 87; Gaps 14;  
QY 1 MKNASNDAPSSDGAAG---LVPEIN-NEVMPLEPVAGSLATPVVGQNIIDPWIRNPF 56  
DB 1 MWMASKDATSSVDGASGAGQLVPEYNASDPLAMDPEVAGSTAVATAGQVNPIDPWIRNPF 60

	147	DVRQLEPFLPMPDIWNFFHYNQGNDBPKLRVLVAMLYTPLRANNSGDDVFTVSC--RVLT	204
QY		:   :   :   :   :   :   :   :   :   :   :   :   :	
Db		:   :   :   :   :   :   :   :   :   :   :   :   :	
	212	DARQVEPVIFAIPDLRSNLHYLMSDTD-TTSLVIMVYNDL-INPYANDTNSSGCIVTVET	269
		:   :   :   :   :   :   :   :   :   :   :   :   :	
QY	205	KPSDPDEFTEFLVPP	218
		:   :   :   :   :   :   :   :   :   :   :   :   :	
Db	270	KPGDPDKFHLKPP	283
		:   :   :   :   :   :   :   :   :   :   :   :   :	
<b>RESULT 3</b>			
US-09-590-020-2			
; Sequence 2, Application US/09590020			
; Patent No. 6355246			
; GENERAL INFORMATION:			
; APPLICANT: Kruger, John M			
; APPLICANT: Maes, Roger K			
; APPLICANT: Vilnis, Aivars			
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE			
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF			
; FILE REFERENCE: MSU 4.1-446			
; CURRENT APPLICATION NUMBER: US/09/590,020			
; CURRENT FILING DATE: 2000-06-08			
; PRIOR APPLICATION NUMBER: 60/138,484			
; PRIOR FILING DATE: 1999-06-10			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: Patentin Ver. 2.1			
; SEQ ID NO 2			
; LENGTH: 623			
; TYPE: PRT			
; ORGANISM: Feline calicivirus			
US-09-590-020-2			
Query Match                    9.2%; Score 265.5; DB 4; Length 623;			
Best Local Similarity       34.0%; Pred.No.1.8e-18;			
Matches     66; Conservative     32; Mismatches   81; Indels   15; Gaps   6			
QY	30	EPVAGASLATPVVGQQNIIDFWIRNNFVQAPAGEFTV---SPRNSGPEITLLDLELGPDLN	86
		:   :   :   :   :   :   :   :   :   :   :   :   :	
Db	100	EPSAQMSAAADMATGKSVDSSEW-----EAFSFTSTVNWSSTSETQGKILFKQSLGPLLN	153
		:   :   :   :   :   :   :   :   :   :   :   :   :	
QY	87	PYLHAARMYNGHAGGEHVQIVLAGNAFTAGKIIFAAIPGGPYENLSPSQITMCPHVII	146
		:   :   :   :   :   :   :   :   :   :   :   :   :	
Db	154	PYLSHAKLYAVWSSGVFRFSIGSGVFGGKLAAIIVPPGI--EPVQSTMLQYPHVLV	211
		:   :   :   :   :   :   :   :   :   :   :   :   :	
QY	147	DVROLEPFLPMPDIWNFFHYNQGNDBPKLRVLVAMLYTPLRANNSGDDVFTVSC--RVLT	204
		:   :   :   :   :   :   :   :   :   :   :   :   :	
Db	212	DARQVEPVIFAIPDLRSNLHYLMSDTD-TTSLVIMVYNDL-INPYANDTNSSGCIVTVET	269
		:   :   :   :   :   :   :   :   :   :   :   :   :	
QY	205	KPSDPDEFTEFLVPP	218
		:   :   :   :   :   :   :   :   :   :   :   :   :	
Db	270	KPGDPDKFHLKPP	283
		:   :   :   :   :   :   :   :   :   :   :   :   :	
<b>RESULT 4</b>			
US-09-590-020-4			
; Sequence 4, Application US/09590020			
; Patent No. 6355246			
; GENERAL INFORMATION:			
; APPLICANT: Kruger, John M			
; APPLICANT: Maes, Roger K			
; APPLICANT: Vilnis, Aivars			
; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE			
; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF			
; FILE REFERENCE: MSU 4.1-446			
; CURRENT APPLICATION NUMBER: US/09/590,020			
; CURRENT FILING DATE: 2000-06-08			
; PRIOR APPLICATION NUMBER: 60/138,484			
; PRIOR FILING DATE: 1999-06-10			
; NUMBER OF SEQ ID NOS: 9			
; SOFTWARE: Patentin Ver. 2.1			
; SEQ ID NO 4			
; LENGTH: 623			
; TYPE: PRT			

57	Y	VOAPAGEFTVSPRNSPGEILLDLEGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTA	116
61	b	VOAQOGEFTISPNNTPGDVLFDLSLGPLNLFLLHSQMYNGWGNRVRIMLAGNAFTA	120
117	Y	GKIIFAAIPGPPYENISFQITCMCPHVIDVROLEPELLPMPDINNFPHYNGNDPKL	176
121	b	GKIIVSCIPPGFGSHNLITIAQATLPHVIAVDTLDPDIEVLEDRNVLFHNDRNQOTM	180
177	Y	RLVAMLYTLRLANNSGDDVFTVSCRVLTKPSDPDEFTFLVPPTVESKTKQALPILKISE	236
181	b	RLVCMLYTLPLRTGGGTGDSFVAVRWMTCPSPDFNFLFLVPDPTVEQKTRPFTLNLPLSS	240
237	Y	MTNSRPPVVDVMTARNENQVVOQNGREVTLDGELGTTTLLAVNICKFKGEVIAKNGD	296
241	b	LSNSRAPLPISSMGISIPDNVQVQFNGRCTLIDGRLVGTTPVLSHVAKIRG---TNGT	297
297	Y	VRSYRMDMETNTDGPIDPDTEDTPGIPSGDFQG-----ILFGVASQNKNEQNPA	349
298	b	V-----INLTLDGTFPHFPEG-PAPIGFPLDGGCDWHINMTQFGHSSQTQYD-----	344
350	Y	RAHEALINTGGDLCTQIISSEIYLTSPNILACTWPOPLQPSGLRGTTILIBSDNGHCHDM	409
345	b	-----VDTPDPTFVPHLGS-----IQANGIGSGNY	369
410	Y	VG-----TSPTTTPTWQ--QWRRCRSGSNCCSSGHRYP--VPVVMNRVTWVLSHKSGFS	460
370	b	VGVLWSIPSPHSFGSQVDLWKIPNYGSSITEATHLAPSVYPPGFGVLFVFMKMPGPG	429
461	Y	TSTRK--LPQLNL-----RWPLRFINPDTGRVLFEARLHKQGITVAHTG--	504
430	b	AYNLPCLLPGEYISHLASEQAPTVGAAALHYVDPDTGRNLGFEKAYPDGFLTVPNGAS	489
505	Y	DNPIVMPNPGYFRFEAWNOFYSLAPVGTGKGR	538
490	db	SGPOQLPINGVVFVSVWSRFPQLKPVGTASSAR	523

RESULT 2

US-09-590-020-6

; Sequence 6, Application US/09590020

; Patent No. 635246

; GENERAL INFORMATION:

; APPLICANT: Kruger, John M

; APPLICANT: Maes, Roger K

; APPLICANT: Vilinis, Aivars

; TITLE OF INVENTION: METHOD FOR THE ISOLATION AND DIAGNOSIS OF FELINE

; TITLE OF INVENTION: CALICIVIRUS AND VACCINES THEREOF

; FILE REFERENCE: MSU 4.1-446

; CURRENT APPLICATION NUMBER: US/09/590,020

; CURRENT FILING DATE: 2000-06-08

; PRIOR APPLICATION NUMBER: 60/138,484

; PRIOR FILING DATE: 1999-06-10

; NUMBER OF SEQ ID NOS: 9

; SOFTWARE: PatentIn Ver. 2.1

; SEQ ID NO 6

; LENGTH: 622

; TYPE: PR

; ORGANISM: Feline calicivirus

US-09-590-020-6

Query Match 9.2%; Score 267.5; DB 4; Length 622;

Best Local Similarity 34.5%; Pred. No. 1.1e-18;

Matches 67; Conservative 31; Mismatches 81; Indels 15; Gaps 6

QY	30	EPVAGASLAPVVGQONTIDWIRNNFVQAPAGEFTV---SPRNSPGEILLDLELGPDLN	86
Db	100	EPSAQSAADMATCKSDVSEW-----EAPFSFTSVNMSTSTQCKILFKQSLGPLN	153
QY	87	PYLALHARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPGPPYENISPSQITMCPHVII	146
Db	154	PYLSHLAKLYVALAGSVEVRSSIGSGVFGKLAIVVPGI--EPVQSTSMLYQPHLVF	211



```

; ORGANISM: Peline calicivirus
US-09-617-594A-2

Query Match      8.7%; Score 250.5; DB 4; Length 669;
Best Local Similarity 25.6%; Pred. No. 7.6e-17;
Matches 93; Conservative 58; Mismatches 141; Indels 71; Gaps 15;

QY 67 SPRNSPGEILLDLLELGDPLNPYLAHLARWYNGHAGGMEYQIVLAGNAFTAGKLIIPAAITPP 126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 180 STSTQGGKILFKQSLGSLLPYLEHLSKLYAVMSGSDVRFSGSGVFGGKLAIAIVVPP 239
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 127 GF-PYENLSPSQITWCCHVIIDVRQLEPFLPLMPDIWNFFHYNQNCDPKRLVLAULTP 185
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 240 GUDPVQSTSMLOY---PHVLFDARQVEPVLFSPDLRSLTYLHLMSTD-TTSLVIMVND 295
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 186 L-----RANNSGDDVFTVSCRVLTKPSDPEFTFLVPPTVESKTKQFALPIKISEMINS 240
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 296 LINPYANDSNSGCGITVE---TKPGDFPKFHLKPPG-----SMLYTHG 336
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 241 RPP---VP-----VDVMYTARNENQVQVP-----QNGRVTLDGELLG-TTPL---LAVNI 283
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 337 SIPSDLIPKSSSLWIGNRYWSDITDFVIRFVQARHDFNQETAGWSTPRFRITITI 396
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 284 CKFKEVIATKN-----GDVRSYRMDMEINTDGTDPIDPTE--DTPG 322
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 397 SESNGSKLGTGVATDIYVPGIDGPDPTTIGELTPAGDYSITNGSGNDIATANAYDSAD 456
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 323 PI-GSPDFQGIILFGVASQRKNQONPATRAHEALINTGGDHLCPQISSSEIVLTSPNLR 381
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 457 VITNTNFRGMYICGALQRAWGDKKISSTAFITAIKEGNLAP---SMTIDMTKIAVYQ 513
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 382 CTN 384
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 514 DTH 516
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
US-09-396-149-5
; Sequence 5, Application US/09396149
; Patent No. 6538176
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Maize Replication Protein A and Use
; FILE REFERENCE: 5718-59
; CURRENT APPLICATION NUMBER: US/09/396,149
; CURRENT FILING DATE: 1999-09-15
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Oryza sativa
US-09-396-149-5

Query Match      3.9%; Score 114; DB 4; Length 630;
Best Local Similarity 20.9%; Pred. No. 0.012; 99; Indels 112; Gaps 16
Matches 67; Conservative 42; Mismatches 42;

QY 123 AIPPG---FPYENLSPSQIT--MCPHVIIDVRQLEPFLPLMPDIWNFF--PHYNGNDP-K 175
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 8 SVTPGAVAFVLENASPDATGVPVPEIVLVQVDLKP-----IGTRFTFLASDGDKKIK 160
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 176 LRLVAMLYTPLRANSGD---DVFTVSCR-----VLTKPSDPDEF----- 212
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 TMLLTQLAPEVRSNGNLGVIRLDYTCNTIGEKQEKVLIITKLEVVFKALDSEIKCEA 120
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 213 -----TFVLVPPTVESKTKQFALPIKISEMNSRFFVPVDVM---YTAENENQVVPQ 262
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 EKQEEKPAILLSPKES-----VVLSEKFTNAP-PLPPVLKPKQEKVKSQSIVNEQ 170
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 263 NGRVTLDELGLLGT---PLLAVNICFKKEVIATK-----NGDVRSYR-----MDMEI 306
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 171 RGNAAAPARLAWTRRVHPLTSL--PYQGNWIAIKRVVTSKGNLRTYKNAGCGVFNVEL 228
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```





; TITLE OF INVENTION: CDNA
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release 1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/465,250
; FILING DATE: 6-JUN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P.
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 36607-E-PCT-US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2206 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-465-250-2
;
; Query Match 3.8%; Score 114; DB 3; Length 2206;
; Best Local Similarity 18.1%; Pred. No. 0.099;
; Matches 136; Conservative 86; Mismatches 237; Indels 294; Gaps 34;
;
; Qy 26 VMLEPV---AGASLATPVV-----GQNIIDP-----WIRNPF 56
; Db 299 ILPLSLDFAQSSVEIPITVTIAPMCSFNGRLNVTAPKFGPLVLTGSGNOYLTSN 358
; Qy 57 VQAPAG--EFTVSPR--NSPGRI-----K 75
; Db 359 HQSPCAIPEFDVTIPDIPGVEKRMELAEIDTWIPLNLESTKENTMDYRVTLSDSADL 418
; Qy 76 ---LLDLELGPLNPLYLAH-----LARMYNGHAGMEVQIVLAGNAFTAKIIPALPPG 127
; Db 419 SQPILCLSLSPAFDPRLSHTMLGVLNYYTHWAGSLKFTFLFCGMMMATGKILVAYAPPG 478
; Qy 128 FPYENLSPQITMCPHVIIIDVQLEPFLPMPDIWNNFFHYNQNDP-----K 175
; Db 479 -AQPTSRKEAMLGTHVIMDLGLOSSCTMVVP--WISNVYRQTQDSFTGGVISMFYQ 535
; Qy 176 LRLVAMLYTPLRANSGDDVFTVSCRVLTKSPDPFEFTL-----VPPTVESKTK 225
; Db 536 TRIVVPLSTPKSMMLG---FVSACN-----DFSVELLADTTHISQALPQGIEDLTS 585
; Qy 226 QFA-----LPIKISEMTNSR--FP-----VPDVMT----- 251
; Db 586 EVAQCALTLSPKQDLSLFDTKASGPAHSKEVPALTAVETGATNPLAPSDTVQTRHVQR 645
; Qy 252 -ARNEN-----QVQPQNGRVTLDGELLGTPLLAVNICPKFGEV-IAGNGDV 297
; Db 646 RSRSESTIESFPARGACVAIIEVDNEQPTTAAQ-----KLFAMWRIITYKDTVQLRKLKF 700
; Qy 298 RSY-RMDME-----ITNTD-----GTPIDPTED----- 319
; Db 701 FYISREFMEFTVVTANFTNANNHALNQVYQIMVPPGAPTPKSWDDTYTWQTSNPSIF 760
; Qy 320 -----TPGPIGSPDQGI-----FGVASQ 339
; Db 761 YTYGAAPARISVP-YVGLANAYSHFYVDGFAKVPVLKTDANDQIGDSLYSAMTVDDFGVLAV 819
;
;
; 340 RNKNEONPATRAHEAINTGGDHL---CPQISSSEIYLTSPNILRCTNPQLPQSGSLRGT 396
; Db 820 RVVDHNPTKVTYSKVIYMKPKHVRVWCPRPRAVPYY-GPGVDYRNNDLPLSEKG---- 874
; Qy 397 ILIRSDNGHCHDMVGTSPPTTTPWPOQWRRCSRSGNSCCS-----SGHRYVPVPMNRVTW 450
; Db 875 -LITYGFGHQNKAITYA-----GYKICNYHLATKEDLQNAVSIMWNRDLL 918
; Qy 451 IVLSHKSFGSTSTRKLPQLNLRWPLIRFINPDTGRVLFPEARLHKQGFITVAHTGDNPIVM 510
; Db 919 VVESKAQGTDSIAR-----CNCNAGVYCESR-----RKYPVSVFGVPTFOYM 961
; Qy 511 PPNGYF--RFEAWNOFYSLAPVGTGKRRRVQ 541
; Db 962 EANDYYPARYQSHMLIGHGFASPDGCGILRCQ 994
;
; RESULT 12
; US-08-986-485-2
; Sequence 2, Application US/08986485
; Patent No. 6046030
; GENERAL INFORMATION:
; APPLICANT: WU, SHUJIAN
; APPLICANT: SWEET, RAYMOND
; APPLICANT: TRUNEH, ALEMSEGED
; TITLE OF INVENTION: A HUMAN LIG-1 HOMOLOG (HLIG-1)
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: RATNER & PRESTIA
; STREET: P.O. BOX 980
; CITY: VALLEY FORGE
; STATE: PA
; COUNTRY: USA
; ZIP: 19482
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/986,485
; FILING DATE: 08-DEC-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/059,448
; FILING DATE: 22-SEP-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: PRESTIA, PAUL F
; REGISTRATION NUMBER: 23,031
; REFERENCE/DOCKET NUMBER: GH-70264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 610-407-0700
; TELEFAX: 610-407-0701
; TELEX: 846169
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1101 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-986-485-2
;
; Query Match 3.8%; Score 109.5; DB 3; Length 1101;
; Best Local Similarity 21.1%; Pred. No. 0.09;
; Matches 120; Conservative 55; Mismatches 193; Indels 201; Gaps 33;
;
; Qy 51 WIRNPFVQAPAG-EFTVSPRNSPGEILLDELGLDPLNRY--LAHLARMYNGHAGMEVQI 107
; Db 538 WKKDNEVLNTADMENFVHQAQGEVW-----EYTTILHLRQVTFGHEG--RYQC 585
; Qy 108 VLA---GNAFTAGKIIFAAIPGFFPYENLSPQITMCPHVI-----IDVRQLEPFL--PM 158

Db 586 VITNFGSTYSHKARLVNVLPSF-----TKTPHDITIRTTTVALECAATGPHN 635  
QY 159 PDI-MNNFFHYNOGND-PKLRVAMLYTPLRANSGDDVFTV-----SC-----200  
Db 636 PQIAWQK-----DGGTDFPAARERRHVP-----DDVFFITDKIDDAGVYSCAQS 685  
QY 201 -----RVLTQPS-----PQFETFLVPPTVESKTKQFALPKIKSEMNTSRFPVPVD 247  
Db 686 AGSISANATLTVLETPSLVPLEDRVVSGETVALQCKATGNPPPRITWFKGDR-PLSLT 744  
QY 248 VMYTARNENOVQPON-----GRVTLG-GEILGT-----TPLLAVNICKEGVEIAKN 294  
Db 745 ERHLLTPDNLQVQNVQVAEDAGRYCEMSNTLGTERRAHSQSLVPAACRKGDTTV---801  
QY 295 G-----DVRSYRMDMEITNDGTPIDPTEDTGPIGSPDFQGI 332  
Db 802 GIFTIAVSSIVLTSVWCIIYQTRKSEESVNTDVTVP--DVESYLSS---QQT 856  
QY 333 LFGVASQRNKNQNPATRAHEAIINTGDLCPQISSSEIYLTSPNI--LRCNTPQ-----386  
Db 857 LSDROETVVRTEGGPOANGH---IESNG--VCPRDAS---HFPEPDTHSVACRQPKLCAG 908  
QY 387 -----PLPQSLRGTILIRSDNGHCHDMV-----CT 412  
Db 909 SAYHKEPWKAMEKAEGTGPHEKHEGGRVYVCS---CNTEVDCYSRQAFHPQPVSRDSA 965  
QY 413 SPTTPTWPO-----QWRRCSRGSCSSGHRYPVPVWNRVTWIVLSHKSQFS 460  
Db 966 QPSAPNGPFGGSDQEHSPHQCSRTAAGSCPECQGSLYP-----SNHDMRL 1012  
QY 461 TSTRKLPQLNL-----RWPLIRFINPDT 483  
Db 1013 TAVKKKPMASLDGKGDSSTWTLARLYHPS 1041  
  
RESULT 13  
US-09-091-219-24  
; Sequence 24, Application US/09091219  
; Patent No. 6171592  
; GENERAL INFORMATION:  
; APPLICANT: STUDDERT, Michael J.  
; APPLICANT: CRABB, Brendan S.  
; APPLICANT: FENG, Li  
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS  
; CURRENT APPLICATION NUMBER: US/09/091,219  
; PRIOR FILING DATE: 1998-10-05  
; EARLIER APPLICATION NUMBER: PCT/AU96/00815  
; EARLIER FILING DATE: 1996-12-18  
; EARLIER APPLICATION NUMBER: AU PN7201  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 2318  
; TYPE: PRT  
; ORGANISM: Foot-and-mouth disease virus  
US-09-091-219-24  
  
Query Match 3.6%; Score 104.5; DB 3; Length 2318;  
Best Local Similarity 22.6%; Pred. No. 1.1;  
Matches 33; Conservative 21; Mismatches 67; Indels 25; Gaps 2;  
  
QY 79 LEIGDPLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKIIFAAPPGFPYENLSPSQI 138  
Db 367 LELPTDHKGVGSLTDSYAYMRGNDVETAVGNQFNGCGLLVAMVPELYSIQKRELYQL 426  
QY 139 TWCPHVVIDVROLEPFLPMPDIWNNFFHYNGNDPKRLVAMLYTPLRANSG-----192  
Db 427 TLFPHQFINPRTNMTAHITVPFGVNRVDQYKVHKP-WTLVVMVAVPLTVNTGAPQIKV 485  
QY 193 -----DDVFTVSC 200  
Db 486 YANIAPTNNVHAGBPPSKGPIPVAC 511  
  
RESULT 15  
US-09-489-039A-9905  
; Sequence 9905, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 9905  
; LENGTH: 596  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9905  
  
Query Match 3.5%; Score 102.5; DB 4; Length 596;  
Best Local Similarity 21.4%; Pred. No. 0.17;  
Matches 112; Conservative 52; Mismatches 221; Indels 139; Gaps 24;  
  
QY 99 HAGGMEVQIVLAGNAFTAGKIIFAAPPGFPYENLSPSQITMCP-----HVLIIDVRLQRP 153  
Db 52 HAGCARVECVIADTCI-AGMAESAACEKFSNQNGVT--ITVPCWCYGSSETIDMDPLRP 109

Db 486 YANIAPTNNVHAGBPPSKGPIPVAC 511  
  
RESULT 14  
US-09-660-541-24  
; Sequence 24, Application US/09660541  
; Patent No. 6531136  
; GENERAL INFORMATION:  
; APPLICANT: STUDDERT, Michael J.  
; APPLICANT: CRABB, Brendan S.  
; APPLICANT: FENG, Li  
; TITLE OF INVENTION: EQUINE RHINOVIRUS 1 PROTEINS  
; FILE REFERENCE: 040268/0151  
; CURRENT APPLICATION NUMBER: US/09/660,541  
; CURRENT FILING DATE: 2000-09-12  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/091,219  
; PRIOR FILING DATE: EARLIER FILING DATE: 1998-10-05  
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: AU PN7201  
; PRIOR FILING DATE: EARLIER FILING DATE: 1995-12-18  
; NUMBER OF SEQ ID NOS: 25  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 24  
; LENGTH: 2318  
; TYPE: PRT  
; ORGANISM: Foot-and-mouth disease virus  
US-09-660-541-24  
  
Query Match 3.6%; Score 104.5; DB 4; Length 2318;  
Best Local Similarity 22.6%; Pred. No. 1.1;  
Matches 33; Conservative 21; Mismatches 67; Indels 25; Gaps 2;  
  
QY 79 LEIGDPLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKIIFAAPPGFPYENLSPSQI 138  
Db 367 LELPTDHKGVGSLTDSYAYMRGNDVETAVGNQFNGCGLLVAMVPELYSIQKRELYQL 426  
QY 139 TWCPHVVIDVROLEPFLPMPDIWNNFFHYNGNDPKRLVAMLYTPLRANSG-----192  
Db 427 TLFPHQFINPRTNMTAHITVPFGVNRVDQYKVHKP-WTLVVMVAVPLTVNTGAPQIKV 485  
QY 193 -----DDVFTVSC 200  
Db 486 YANIAPTNNVHAGBPPSKGPIPVAC 511  
  
RESULT 15  
US-09-489-039A-9905  
; Sequence 9905, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 9905  
; LENGTH: 596  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9905  
  
Query Match 3.5%; Score 102.5; DB 4; Length 596;  
Best Local Similarity 21.4%; Pred. No. 0.17;  
Matches 112; Conservative 52; Mismatches 221; Indels 139; Gaps 24;  
  
QY 99 HAGGMEVQIVLAGNAFTAGKIIFAAPPGFPYENLSPSQITMCP-----HVLIIDVRLQRP 153  
Db 52 HAGCARVECVIADTCI-AGMAESAACEKFSNQNGVT--ITVPCWCYGSSETIDMDPLRP 109

QY 154 FLLPMPDIWNPFHYNQGNPKRLVAMLYTPLRANNS-----GDDVFTVSCRVL 204  
Db 110 -----KAIWG-----ENGTERPG---AVYLAALAAHSQKIPAFSIYGHV----- 148  
QY 205 KPSPDFEFTFLVPPPTVESKTKQFALPILKISEMTNSRF-----PVPVDVMTARNEN----- 256  
Db 149 ---QDADDT-SIPADVEEKLRLPFAAGLAVASMKGSYLSLGGVSMGIAGSIVDHNPFES 204  
QY 257 -----QVVQPQNGRVTLDGELLGTTPLIAVNICFKGEVIAKNGDVRSYRMD----- 303  
Db 205 WLGKMQVQAVDMTELRRRIDOKIYDETELEMAWADKHFYGEDQNAEQYKRNETQSRV 264  
QY 304 -----MEITNTDGTIDPTDTPGPI-GSPDFQGIILFGVASQORNKNEQNPAHRAHAI 356  
Db 265 LKESLIMAMCIRDMMQGNPKLAEKGLVEESLGYNAIAGFQQRHWTQYFNGDTAEALL 324  
QY 357 N-----TGGDHL-----CPQISSEIYLTSPNLRCTNPQPL 388  
Db 325 NSSFDMNGVREPFVATENDSLNGVAMLMGHQLTGTAQVFADVRTYWSPDVERVTGQPL 384  
QY 389 PQSGLRGTI-LIRSD-----NGHC--HDMVGTSPPTPTW---PQWRRCSRGNCCSSGHR 438  
Db 385 TGRAEHGIIHLINSQSAALDGSQQORDAQCNPMTKPHWEIEQSEADACLAATEWCFAIHE 444  
QY 439 Y-----PVPVMMNRVTWI-----VLSHKSGFSTSTRKL--POLNLR-----W 473  
Db 445 YFRGGGFSRFLTEGGVPFTMTTRVNIKGLGPVLQIAGWSVALPKAMHDQDARTNSTW 504  
QY 474 PLIRFINPDTGRVLFEARLH-----KQGFITVAHTGDNPIVM 510  
Db 505 PTTWFAPRLTGKGFSDVYSVMANWGANHGVLTIHVGADFITL 548

Search completed: June 1, 2004, 13:58:11  
Job time : 14.8896 secs

GenCore version 5.1.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:46:07 ; Search time 34.5841 Seconds

(without alignments)  
4368.312 Million cell updates/sec

Title: US-09-926-799-11

Perfect score: 2894

Sequence: 1 MKMASNDAPSDGAGLVP.....VNQFSLAPVGTGKRRRVQ 541

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1151071 seqs, 279249464 residues

Total number of hits satisfying chosen parameters: 1151071

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCT05\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09A\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09B\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09C\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1008.5	34.8	530	14	US-10-314-739-3
2	264	9.1	668	14	US-10-209-507-4
3	262.5	9.1	547	12	US-10-670-695-4
4	262.5	9.1	671	12	US-10-670-695-2
5	250.5	8.7	669	14	US-10-209-507-2
6	118.5	4.1	1147	15	US-10-327-481A-38
7	114	3.9	630	14	US-10-372-686-5
8	114	3.9	630	14	US-10-371-558-5
9	114	3.9	630	14	US-10-375-553-5
10	114	3.9	630	16	US-10-372-553-5
11	108.5	3.7	861	12	US-10-282-122A-66008
12	108	3.7	3696	16	US-10-229-148B-4
13	107	3.7	607	12	US-10-258-951-74
14	106.5	3.7	767	12	US-10-296-115-1211
15	106.5	3.7	861	12	US-10-282-122A-65330

16	106.5	3.7	1077	15	US-10-093-463-136	Sequence 136, Appl
17	103	3.6	578	9	US-09-924-358-39	Sequence 39, Appl
18	103	3.6	578	15	US-10-410-764-39	Sequence 39, Appl
19	102.5	3.5	656	14	US-10-128-714-3343	Sequence 3343, Ap
20	102.5	3.5	1049	12	US-10-211-462-6	Sequence 6, Appli
21	102.5	3.5	1049	12	US-10-170-385-379	Sequence 379, App
22	102.5	3.5	1049	15	US-10-236-031B-52	Sequence 52, Appl
23	102.5	3.5	1049	15	US-10-295-027-4	Sequence 4, Appli
24	102.5	3.5	1049	15	US-10-295-027-726	Sequence 726, App
25	102.5	3.5	1049	15	US-10-295-027-1166	Sequence 1166, Ap
26	102.5	3.5	1049	15	US-10-301-330-2	Sequence 2, Appli
27	102.5	3.5	1049	15	US-10-188-832-2	Sequence 2, Appli
28	102.5	3.5	1063	12	US-10-425-114-39245	Sequence 39245, A
29	101.5	3.5	639	12	US-10-424-599-144381	Sequence 144381, A
30	100.5	3.5	723	12	US-10-425-114-38957	Sequence 38957, A
31	100.5	3.5	837	9	US-09-738-626-6913	Sequence 6913, Ap
32	99.5	3.4	745	14	US-10-128-714-8343	Sequence 8343, Ap
33	98.5	3.4	481	15	US-10-389-566-826	Sequence 826, App
34	98.5	3.4	1936	12	US-10-152-886-1	Sequence 1, Appl
35	98.5	3.4	2781	12	US-10-263-929-122	Sequence 122, App
36	98	3.4	564	14	US-10-156-761-8951	Sequence 8951, Ap
37	98	3.4	3931	15	US-10-120-801-18	Sequence 18, Appl
38	97.5	3.4	4952	15	US-10-051-874-56	Sequence 56, Appl
39	97.5	3.4	5008	15	US-10-051-874-166	Sequence 166, App
40	97.5	3.4	5159	15	US-10-085-198-112	Sequence 112, App
41	97	3.4	865	12	US-10-282-122A-77101	Sequence 77101, A
42	97	3.4	2515	12	US-10-042-865-53	Sequence 53, Appl
43	97	3.4	2515	12	US-10-029-020-53	Sequence 53, Appl
44	97	3.4	6310	12	US-10-282-122A-67793	Sequence 67793, A
45	96.5	3.3	262	14	US-10-059-909-10	Sequence 10, Appl

ALIGNMENTS

RESULT 1  
US-10-314-739-3  
; Sequence 3, Application US/10314739  
; Publication No. US20030129588A1  
; GENERAL INFORMATION:  
; APPLICANT: Bates, Mary K  
; Jiang, Xi  
; Graham, David Y  
; TITLE OF INVENTION: Methods and Reagents to Detect and  
; Characterize No. US20030129588A1walk and Related Viruses  
; NUMBER OF SEQUENCES: 4  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pubbright & Jaworski L.L.P.  
; STREET: 801 Pennsylvania Ave., N.W.  
; CITY: Washington, D.C.  
; STATE: <Unknown>  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA: US/10/314,739  
; APPLICATION NUMBER: US/10/314,739  
; FILING DATE: 09-Dec-2002  
; CLASSIFICATION: <Unknown>  
; PRIORITY INFORMATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/486,049  
; FILING DATE: June 7, 1995  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Davis, Peter  
; REGISTRATION NUMBER: 36,119  
; REFERENCE/DOCKET NUMBER: 311.023  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202-662-0200  
; TELEFAX: 202-662-4643  
; TELEX: <Unknown>

```
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
;   LENGTH: 530 amino acids
;   TYPE: amino acid
;   TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-10-314-739-3

Query Match          34.8%; Score 1008.5; DB 14; Length 530;
Best Local Similarity 40.9%; Pred. No. 6.2e-87;
Matches 235; Conservative 72; Mismatches 180; Indels 87; Gaps 14;

QY 1 MKWASNDAAAPSDGAG---LVEPIN-NEVMELEPVAGASLATPVVQGNIIIDPIRNWF 56
DB 1 MMASKDATSSVDGASGAGLVPEVNASDPLAMDVPAGSSTAVATAGVNPIDPIWNNF 60
QY 57 VQAPAGEFTVSPRNSGEILLDLGLPDINPVLAHLARMYNGHAGMEVQIVLAGNAFTA 116
DB 61 VQAPQGEFTISPNNTPGDLVFLDSLGLPHLPFLHLSQMYNGWGNMVRIMLAGNAFTA 120
QY 117 GKLIIFAAIPGFPYENLSPSQITMCPHVIIDVROLEPFLPMPDIWNFFHYNQGNPKL 176
DB 121 GKLIIVCIPGFGSHNTIAQATLFPHVIADVRLDPIEVPLEDVRNVLFNHNDNRNOQM 180
QY 177 RLVMALYTPLRANSGDDVFTVSCRVLTKPSDPDEFPLVPPTVESKTKQFALPILKISE 236
DB 181 RLVCMLYTPLRITGGTGDSFVAGRVMTCPSPDFNLFPLVPPTVEQKTRPFTLPLPLSS 240
QY 237 MTNSRPVPVDMVTARNENQVQPNQGRVTLDELGLLGTTPLLAVNICFKGKGVIAKNGD 296
DB 241 LNSRAPLPISMGISPDNVQSVQFNQGRCTLDRLVGTTPVSLSHVAKIRG---TSNGT 297
QY 297 VRSYRDMETINTDGTPIDPTEDTPGIGSPDFQG-----LFGVAGSQRNKNQNPAT 349
DB 298 V-----INLTEDGTFPHFEG-PAPIGFDFDLGGCDWHINMTQFGHSSQTYD----- 344
QY 350 RAHEALINTGDHLCPQISSEIYLSPTNLRTCNQPLQPSGLRGTLILRSNCHDM 409
DB 345 -----VDITDPTFVPLHGS-----IQANGISGNV 369
QY 410 VG-----TSPTTPTWPO--QWRRCRSGNSCCSSGHRYP--VPVVMNRVTWIVLSHSGFS 460
DB 370 VGVLSWSPSHSGSGQVDLWKIPNYGSSITEATHLAPSVYPGFGVLFVFFSKMPGPG 429
QY 461 TSTRK--LPQLNL-----RWPLIRFNPDTRVLFEARLHKQGFITVAHTG-- 504
DB 430 AYNLPCLLPQEIYISHLASEQAPTVEAALLHYVDPDTRNLGEFKAYPDGFLTCVPNGAS 489
QY 505 DNPIWPNPGVFRFEMVNOFYSVLPVGTGKGR 538
DB 490 SGPOQLPINGVFVFSVWSREYQLKPVGTASSAR 523

RESULT 2
US-10-209-507-4
; Sequence 4, Application US/10209507
; Publication No. US20030109033A1
; GENERAL INFORMATION:
; APPLICANT: Audonnet, et al.
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT V
; FILE REFERENCE: 454313-3151-2
; CURRENT FILING DATE: 2002-11-12
; PRIOR FILING DATE: 09/617,594
; PRIOR FILING DATE: 2000-07-14
; PRIOR FILING DATE: 60/193,332
; PRIOR FILING DATE: 2000-03-30
; PRIOR FILING DATE: France 00 01761
; PRIOR FILING DATE: 2000-02-11
; PRIOR FILING DATE: France 99 09421
; PRIOR FILING DATE: 1999-07-16
; NUMBER OF SEQ ID NOS: 26

Query Match          9.1%; Score 262.5; DB 12; Length 547;
Best Local Similarity 25.3%; Pred. No. 8.9e-16;
Matches 113; Conservative 59; Mismatches 167; Indels 107; Gaps 20;

QY 30 EPVAGASLATPVVQGNIIIDPIRNQVQAPAGEFTV---SPRNSPGEILLDLGLPDIN 86
DB 21 EPSAQMSTAADWATKGSVDSEW-----EAFPSFHTSVNMWSTSETQGGKILPKQSLGPLLN 74
QY 87 PYLAHLARMYNGHAGMEVQIVLAGNAFTAIIIPGPF-PYENLSPSQITMCPHVI 145
DB 75 PYLEHLAKLYVAVWSGSIEVRFSISGSGVFGKGLAAIVVPPGVDPVQSTSMLOY---PHVL 131
QY 146 IDVRQLEPFLPMPDIWNFFHYNQGNPKLRLVAMLYTPLRANSGDDVFTVSC--RVL 203
DB 132 FQARQVEPVI FCLPDLRLSTLYHLMSTD-TTSLVIMVNDL-INPYANDANSGCCIVTVE 189
QY 204 TKPSPDPFEFTFLVPPTVESKTKQFALPILKISEMNSRFP---VP-----VDVMYTARN 254
DB 190 TKEGPDFKFLHAKPPG-----SMLTHGSIPLIPKTPKSLWIGNRYWSDI 234
QY 255 ENQVQVQF---QNGRVTLDGELLG-TTPL---LAVNICFKGKGVIAKNGDVRYSRM---- 302
DB 235 TDFVIRPFVQANRHFDFNQETAGWSTPRFRPISVTITEQNGAKLG-IGVATDIYVPGIP 293
QY 303 -----DMEITNTDGTPIDPTD--DTPGPI-GSPDPFGILFVAGSQRNKN 343
```

Db 294 DCGPDTTIRGELIPAGDYAITNGTNDITATGYDTADIIKNTNPRGMYICGSLQRAWG 353  
Qy 344 EQNPATRAHEALINTGCD-----HLCPOISSSEIYL-----TSPNLL----- 380  
Db 354 DKKISNTAFTTATLDGNNKINPCNTIDQSKIIVFQDNHVGKKAQTSDDTLALLGYTG 413  
Qy 381 -----RCTNPQPLPQSGLRG 395  
Db 414 ICEQAIGSDRDRVVRISTILPETGARG 439

## RESULT 4

US-10-670-695-2  
; Sequence 2, Application US/10670695  
; Publication No. US20040058316A1  
; GENERAL INFORMATION:  
; APPLICANT: Jensen, Wayne A.  
; APPLICANT: Lappin, Michael R.  
; APPLICANT: Rosen, David K.  
; APPLICANT: Andrews, Janet S.  
; TITLE OF INVENTION: USE OF RECOMBINANT ANTIGENS TO DETERMINE THE IMMUNE  
; FILE REFERENCE: DI-9-1  
; CURRENT APPLICATION NUMBER: US/10/670,695  
; CURRENT FILING DATE: 2003-09-25  
; PRIOR APPLICATION NUMBER: 09/521,738  
; PRIOR FILING DATE: 2000-03-09  
; NUMBER OF SEQ ID NOS: 36  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 671  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-10-670-695-2

Query Match 9.1%; Score 262.5; DB 12; Length 671;  
Best Local Similarity 25.3%; Pred. No. 1.2e-15;  
Matches 113; Conservative 59; Mismatches 167; Indels 107; Gaps 20;  
Qy 30 EPVAGASLATPVGOQNIIDPIRNNFVOAPAGEFTV---SPRNSPGEILLDLGLPDLN 86  
Db 145 EFSAQMSTAADMATGKSVDSW-----EAFSFTSVNWSSETQGKILFKQSLGPLN 198  
Qy 87 PYLAHLARMYNGHAGMEYQIVLAGNAFTAGKIIFAAIPPGF-PYENLSPSQITMCPHYI 145  
Db 199 PYLEHLAKLYVAWSGSIEVRSISGSGVFGKLAIVVPPGVPQSTSMLOY---PHVL 255  
Qy 146 IDVRQLEPFLPMPDIWNFFHYNQNDPKLRLVAMLYTPTLRANSGDDVFTVSC--RVL 203  
Db 256 FQARQVEPVIFCLPDLRSTLYHLSMTD-TTSLVIMVYNDL-INPYANDANSGCCITVVE 313  
Qy 204 TKPSDPDFEFTFVPPVTSKTKQFALPILKISEMTNSREP---VP-----VDVMYTARN 254  
Db 314 TKPGDPFKFHLKPPG-----SMLTHGSIPSLIPKTSLSLWNGRYWSDI 358  
Qy 255 ENQVQVQ---QNGRVTLDGELLG-TTPL---LAVNICFKGEVIKAGNDVRSYRM---- 302  
Db 359 TDFVIRPFVQANRHFDFNQETAGSTPRFRISVITTEQNGAKLG-IGVATDIYVPGIP 417  
Qy 303 -----DMEITNDGTPIDPTE--DTPGPI-GSPDFQGLIFGVASQNRKN 343  
Db 418 DCGPDTTIRGELIPAGDYAITNGTNDITATGYDTADIIKNTNPRGMYICGSLQRAWG 477  
Qy 344 EQNPATRAHEALINTGCD-----HLCPOISSSEIYL-----TSPNLL----- 380  
Db 478 DKKISNTAFTTATLDGNNKINPCNTIDQSKIIVFQDNHVGKKAQTSDDTLALLGYTG 537  
Qy 381 -----RCTNPQPLPQSGLRG 395  
Db 538 ICEQAIGSDRDRVVRISTILPETGARG 563

## RESULT 5

US-10-209-507-2  
; Sequence 2, Application US/10209507  
; Publication No. US20030109033A1  
; GENERAL INFORMATION:  
; APPLICANT: Audonnet, et al.  
; TITLE OF INVENTION: FELINE CALICIVIRUS GENES AND VACCINES IN PARTICULAR RECOMBINANT  
; FILE REFERENCE: 454313-3151.2  
; CURRENT APPLICATION NUMBER: US/10/209,507  
; CURRENT FILING DATE: 2002-11-12  
; PRIOR APPLICATION NUMBER: 09/617,594  
; PRIOR FILING DATE: 2000-07-14  
; PRIOR APPLICATION NUMBER: 60/193,332  
; PRIOR FILING DATE: 2000-03-30  
; PRIOR APPLICATION NUMBER: France 00 01761  
; PRIOR FILING DATE: 2000-02-11  
; PRIOR APPLICATION NUMBER: France 99 09421  
; PRIOR FILING DATE: 1999-07-16  
; NUMBER OF SEQ ID NOS: 26  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 669  
; TYPE: PRT  
; ORGANISM: Feline calicivirus  
US-10-209-507-2

Query Match 8.7%; Score 250.5; DB 14; Length 669;  
Best Local Similarity 25.6%; Pred. No. 1.7e-14;  
Matches 93; Conservative 58; Mismatches 141; Indels 71; Gaps 15;  
Qy 67 SPRNSPGEILLDLGLPDLNPLYLAHLARMYNGHAGMEYQIVLAGNAFTAGKIIFAAIPPP 126  
Db 180 STSETQGKILFKQSLGPLNPLYLEHLSKLYVAWSGSDVRSISGSGVFGKLAIVVPP 239  
Qy 127 GF-PYENLSPSQITMCPHYIIDVRQLEPFLPMPDIWNFFHYNQNDPKLRLVAMLYTP 185  
Db 240 GVDVPQSTSMLOY---PHVLFQARQVEPVISIPDLRSTLYHLSMTD-TTSLVIMVYND 295  
Qy 186 L-----RANSGDDVFTVSCRVLTKPSPDFEFTFVPPVTSKTKQFALPILKISEMTNS 240  
Db 296 LINPYANDANSNGCIVTVE---TKPGDPFKFHLKPPG-----SMLTHG 336  
Qy 241 RPP---VP-----VDVMYTARNENQVQVQ---QNGRVTLDGELLG-TTTP---LAVNI 283  
Db 337 SIPSLIPKSSLSLWNGRYWSDITDFVIRPFVQANRHFDFNQETAGSTPRFRITITI 396  
Qy 284 CKFKCEVIATKN-----GDVRSYRMDMEITNDGTPIDPTE--DTPG 322  
Db 397 SESNGSKLGTGVATDIYVPGIPDGWPDITIGELTPAGDYSITNGSGNDIATANAYDSAD 456  
Qy 323 PI-GSPDFQGLIFGVASQNRKNEQNPAATRAHEAIIINTGCDHLCPOISSSEIYLTSPNLR 381  
Db 457 VITNTNFRGMVYICGALQORAGDKKISSATFITTAIKSGNTLKP---SNTIDMTKIAVVQ 513  
Qy 382 CTN 384  
Db 514 DTH 516

## RESULT 6

US-10-327-481A-38  
; Sequence 38, Application US/10327481A  
; Publication No. US20040001864A1  
; GENERAL INFORMATION:  
; APPLICANT: King, Andrew M.Q.  
; APPLICANT: Burman, Alison J.  
; APPLICANT: Audonnet, Jean-Christophe F.  
; APPLICANT: Lombard, Michel F.A.  
; TITLE OF INVENTION: Vaccine against Foot-and-Mouth Disease  
; FILE REFERENCE: 454313-3178  
; CURRENT APPLICATION NUMBER: US/10/327,481A  
; CURRENT FILING DATE: 2002-12-20  
; PRIOR APPLICATION NUMBER: PCT/FR01/02042

Wed Jun 2 09:13:25 2004

```

; PRIOR FILING DATE: 2001-06-27
; PRIOR APPLICATION NUMBER: FR 0008437
; PRIOR FILING DATE: 2000-06-29
; NUMBER OF SEQ ID NOS: 41
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 38
; LENGTH: 1147
; TYPE: PRT
; ORGANISM: Foot-and-mouth disease virus
US-10-327-481A-38

Query Match
Best Local Similarity 4.1%; Score 118.5; DB 15; Length 1147;
Matches 77; Conservative 55; Mismatches 156; Indels 137; Gaps 13;

QY 32 VAG-----ASLAPVVGQNIIDPWIRNFFVQAPAGETVSPRNSPGEILDLBLGPDLPNPY 88
DB 129 VAGPNTSGLETRVVQAEFFKKFLPDWTTDKPFGYLT-----KLELPTDHHGV 176

QY 89 LAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPGFPYENLSPSQITMCPHVILDV 148
DB 177 FGHLDVSYAYMRNGWDVEVSAVGNQFNCGCLLVAMVPEWKAFTREKYQLTLFPHQFISP 236

QY 149 RQLEPFLPMPDIWNNFEHYNQNDPKLRLVAMLYTPLRANN----- 190
DB 237 RTNMTAHTITVYLGVNRYDOYKXKHP-WTLVVMVLSPLTVSNTAAPQIKVYANIAPTYVH 295

QY 191 -----SGDDVFTVSC-----RVLTKPSPDF--EFTFLVPPTVES 222
DB 296 VAGELPSKEGIFPVACADGGLVTTDPKTADPVYGVKNYPPKNTYNGRFTNLLDAEAC 355

QY 223 KT-----KQFALPILKISEMTNSRFPVVDVMTARNEN-----QVVPQNGRVT 267
DB 356 PTLRFDDGKPYV-----VTRADDTRLAKFDVSLAAKMSNTYLSGLIAQYTYQYSGTIN 410

QY 268 LDGELLGTTPLLAVNICKERKEVIKNGDVRYSYRMDMEITNDGTIDPTDPTDPTGPIGSP 327
DB 411 LHMFTGSTDSKARYWAY-----IPGVEITP-----P 438

QY 328 DFOGILFGVASQRNKNQNPATRAH-----EALINTGGDHLCPQISSSEIYLTSPNLR 381
DB 439 D-----TPEEAHCIIHAESWDGLNSKFTFSIPYVSAADYAYTASDTAE 481

QY 382 CTNPQ 386
DB 482 TTNVQ 486

RESULT 7
US-10-372-686-5
; Sequence 5, Application US/10372686
; Publication No. US20030159185A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Replication Protein A and Use
; FILE REFERENCE: 0899D
; CURRENT APPLICATION NUMBER: US/10/372,686
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/396,149
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/123,896
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/100,690
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-372-686-5

Query Match
Best Local Similarity 3.9%; Score 114; DB 14; Length 630;
Matches 67; Conservative 42; Mismatches 99; Indels 112; Gaps 16;

QY 123 AIPPG---EPYENLSPSQIT--MCPHVIDVRQLEPFLFLLPMDIWNPF--FHYNQNDP-K 175
DB 8 SVTPGAVAFVLENASPDAAATGVPVEIVLQVVDLKP-----IGTRFTFLASDGKDKIK 60

QY 176 LRLVAMLYTPLRANNSGD-----DVFTVSCR-----VLTKPSPDFEF----- 212
DB 61 TMLLTQLAPEVRSGNIQNLGVIRVLDYTCNTIGEKQKVLIIITKLEVVFKALDSEIKCEA 120

QY 213 -----TFLVPPTVESKTKQFALPILKISEMTNSRFPVVDVM---YTARNENQVVPQ 262
DB 121 EKQEEKPAILLSPKEES-----VVLSKPTNAP-PLPPVVLKPKQEVKSASQIVNEQ 170

QY 263 NGRVTLDGELLGTT---PLLAVNICKFKGEVIKAK-----NGDVRSYR-----MDMEI 306
DB 171 RGNAPAAPARLAMTRRVHPLISLN--PYQGNWIIKVRVTSKGNLRTYKNARGEGCVFNVEL 228

QY 307 TNDGTIDPTDPTDPTGPIGSPDFQGIILFGVASQRNKNQNPATRAHEAIINTGGDHLCPQ 366
DB 229 TDVGTQIQAI-----MFNEAAKKTYPM 251

QY 367 ISSSEIYLTSPNLRCTNPQ 386
DB 252 FELGKVVYISKSLRVANKQ 271

RESULT 8
US-10-371-558-5
; Sequence 5, Application US/10371558
; Publication No. US20030163840A1
; GENERAL INFORMATION:
; APPLICANT: Mahajan, Pramod B.
; TITLE OF INVENTION: Replication Protein A and Use
; FILE REFERENCE: 0899D2
; CURRENT APPLICATION NUMBER: US/10/371,558
; CURRENT FILING DATE: 2003-02-21
; PRIOR APPLICATION NUMBER: US 09/396,149
; PRIOR FILING DATE: 1999-12-15
; PRIOR APPLICATION NUMBER: US 60/123,896
; PRIOR FILING DATE: 1999-03-11
; PRIOR APPLICATION NUMBER: US 60/100,690
; NUMBER OF SEQ ID NOS: 22
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 5
; LENGTH: 630
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-371-558-5

Query Match
Best Local Similarity 3.9%; Score 114; DB 14; Length 630;
Matches 67; Conservative 42; Mismatches 99; Indels 112; Gaps 16;

QY 123 AIPPG---EPYENLSPSQIT--MCPHVIDVRQLEPFLFLLPMDIWNPF--FHYNQNDP-K 175
DB 8 SVTPGAVAFVLENASPDAAATGVPVEIVLQVVDLKP-----IGTRFTFLASDGKDKIK 60

QY 176 LRLVAMLYTPLRANNSGD-----DVFTVSCR-----VLTKPSPDFEF----- 212
DB 61 TMLLTQLAPEVRSGNIQNLGVIRVLDYTCNTIGEKQKVLIIITKLEVVFKALDSEIKCEA 120

QY 213 -----TFLVPPTVESKTKQFALPILKISEMTNSRFPVVDVM---YTARNENQVVPQ 262
DB 121 EKQEEKPAILLSPKEES-----VVLSKPTNAP-PLPPVVLKPKQEVKSASQIVNEQ 170

QY 263 NGRVTLDGELLGTT---PLLAVNICKFKGEVIKAK-----NGDVRSYR-----MDMEI 306
DB 171 RGNAPAAPARLAMTRRVHPLISLN--PYQGNWIIKVRVTSKGNLRTYKNARGEGCVFNVEL 228

QY 307 TNDGTIDPTDPTDPTGPIGSPDFQGIILFGVASQRNKNQNPATRAHEAIINTGGDHLCPQ 366
```



Db 229 TDVDTGTQIAT-----MFNEAAKKFYPM 251  
QY 367 ISSSEIYLTSPNLRCTNPQ 386  
Db 252 FELGKVVYISKGLRVANKQ 271

RESULT 9

US-10-375-553-5  
; Sequence 5, Application US/10375553  
; Publication No. US20030174240A1  
; GENERAL INFORMATION:  
; APPLICANT: Mahajan, Pramod B.  
; TITLE OF INVENTION: Replication Protein A and Use  
; FILE REFERENCE: 0899D3  
; CURRENT APPLICATION NUMBER: US/10/375,553  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: US 09/396,149  
; PRIOR FILING DATE: 1999-12-15  
; PRIOR APPLICATION NUMBER: US 60/123,896  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 60/100,690  
; PRIOR FILING DATE: 1998-09-17  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 630  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-375-553-5

Query Match 3.9%; Score 114; DB 14; Length 630;

Best Local Similarity 20.9%; Pred. No. 0.16;  
Matches 67; Conservative 42; Mismatches 99; Indels 112; Gaps 16;

QY 123 AIPPG---PPYENLSPSQIT--MCPHVIVDVRQLPFLPMPDIWNNF-FHYNQNDP-K 175  
Db 8 SVTFCGAVAFVLENASPDATGVPVPEIVLQVVDLKP-----IGTRFTFLASDGKDKIK 60  
QY 176 LRLVAMLYTPLRANNSGD---DVFVTSR-----VLTKPSDPDEF----- 212  
Db 61 TWMLTQLAPEVRSGNIQNLGVIRVLDYTCNTIGEKQKVLIITKLEVVFKALDSBIKCEA 120  
QY 213 -----TFLVPPTVESKTKQFALPILKISEMTNSRFPVPVDVM---YTARNENQVVPQ 262  
Db 121 EKQEEKPAILLSPKEES-----VVLKSKPTNAP-PLPPVVLKPKQEVKSASQIVNEQ 170  
QY 263 NGRVTLDBELLGTT---PLLAVNICKFKGEVIKAK-----NGDVSRYR-----MDMEI 306  
Db 171 RGNAAAPAAARLAMTRRVHPLISLN--PYQGNWIKVVTSGNLRVTKYKARGGCGVFNVEL 228  
QY 307 TWTGTPIDPTEDTGPISGPDFOGILFVGASQRNKEQNAPTRAHEAINTGGHLCQP 366  
Db 229 TDVDTGTQIAT-----MFNEAAKKFYPM 251  
QY 367 ISSSEIYLTSPNLRCTNPQ 386  
Db 252 FELGKVVYISKGLRVANKQ 271

RESULT 10

US-10-372-553-5  
; Sequence 5, Application US/10372553  
; Publication No. US20040098769A1  
; GENERAL INFORMATION:  
; APPLICANT: Mahajan, Pramod B.  
; TITLE OF INVENTION: Replication Protein A and Use  
; FILE REFERENCE: 0899D3  
; CURRENT APPLICATION NUMBER: US/10/372,553  
; CURRENT FILING DATE: 2003-02-21  
; PRIOR APPLICATION NUMBER: US 09/396,149  
; PRIOR FILING DATE: 1999-12-15

; PRIOR APPLICATION NUMBER: US 60/123,896  
; PRIOR FILING DATE: 1999-03-11  
; PRIOR APPLICATION NUMBER: US 60/100,690  
; PRIOR FILING DATE: 1998-09-17  
; NUMBER OF SEQ ID NOS: 22  
; SOFTWARE: FastSeq for Windows Version 3.0  
; SEQ ID NO 5  
; LENGTH: 630  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
US-10-372-553-5

Query Match 3.9%; Score 114; DB 16; Length 630;

Best Local Similarity 20.9%; Pred. No. 0.16;  
Matches 67; Conservative 42; Mismatches 99; Indels 112; Gaps 16;

QY 123 AIPPG---PPYENLSPSQIT--MCPHVIVDVRQLPFLPMPDIWNNF-FHYNQNDP-K 175  
Db 8 SVTFCGAVAFVLENASPDATGVPVPEIVLQVVDLKP-----IGTRFTFLASDGKDKIK 60  
QY 176 LRLVAMLYTPLRANNSGD---DVFVTSR-----VLTKPSDPDEF----- 212  
Db 61 TWMLTQLAPEVRSGNIQNLGVIRVLDYTCNTIGEKQKVLIITKLEVVFKALDSBIKCEA 120  
QY 213 -----TFLVPPTVESKTKQFALPILKISEMTNSRFPVPVDVM---YTARNENQVVPQ 262  
Db 121 EKQEEKPAILLSPKEES-----VVLKSKPTNAP-PLPPVVLKPKQEVKSASQIVNEQ 170  
QY 263 NGRVTLDBELLGTT---PLLAVNICKFKGEVIKAK-----NGDVSRYR-----MDMEI 306  
Db 171 RGNAAAPAAARLAMTRRVHPLISLN--PYQGNWIKVVTSGNLRVTKYKARGGCGVFNVEL 228  
QY 307 TWTGTPIDPTEDTGPISGPDFOGILFVGASQRNKEQNAPTRAHEAINTGGHLCQP 366  
Db 229 TDVDTGTQIAT-----MFNEAAKKFYPM 251  
QY 367 ISSSEIYLTSPNLRCTNPQ 386  
Db 252 FELGKVVYISKGLRVANKQ 271

RESULT 11

US-10-282-122A-66008  
; Sequence 66008, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Lianguo  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELITRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578

```

; TITLE OF INVENTION: Midecamycin biosynthetic genes
; FILE REFERENCE: 138451 US
; CURRENT APPLICATION NUMBER: US/10/229,148B
; PRIOR FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 210516/2002
; PRIOR FILING DATE: 2002-07-19
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 4
; LENGTH: 3696
; TYPE: PRT
; ORGANISM: Streptomyces mycarofaciens
; US-10-229-148B-4

Query Match      3.7%; Score 108; DB 16; Length 3696;
Best Local Similarity 19.7%; Pred. No. 8.6;
Matches 147; Conservative 72; Mismatches 223; Indels 306; Gaps 37;

QY      8 AAPSSDGAAG-----LVPEINNEVMPLEPVAGASLATPVVQQNI-IDPWIRNNFVOAPA 61
Db      1930 AGISSFGISGNTNVHIVVEE-----PPEPTAPESLWPDAAADGDVWSEWMRE--VTVP- 1980

QY      62 GEFTVSPRNSPG-----EILDLLELGPDLNP-----YLAHLARMYNGHAG----- 101
Db      1981 --LMSAHNEAALCDQARRLRADLLAHPHADVGLSLITRTRFEHRAAVGENTEL 2038

QY      102 -----GMEVOIVLAGNAFTAGKIIFAAIPGPPY-----ENLSFSQ 137
Db      2039 IAALEDLLLEGPHPLVMRGTAGTADQVVFVPGQSSQWAEWGDGLFERSVFRETAHACD 2098

QY      138 ITMCPHV---TIDVRQLEP-----FLLPMPDIWNPF-----FHVNQ 170
Db      2099 AALREYDWSVLDVLRREPDPASLDRVDVQVPLFTMMVSLAATWRSIGVPEPAVVGHSQ 2158

QY      171 GN-----DPKRLVAMLYTLR-----ANNSGDDVFTVSCRVLTKPSPDFE 211
Db      2159 GEIAAAHVAGLSLDDAARIVA-----LRSQAWLQACKGGWVAVTHSEREL---RPLR 2210

QY      212 F-----TFVLVPTVESKTKQFALPILKISEMTNSRFPV-PVDVMYTARNENQVVOQP 262
Db      2211 FWGDRLAVAANVSPETCAVAGDPDALAEL-VAELASQGVPARPIPGVDTAGHSQV--- 2265

QY      263 NGRVTLDELGLGTTPLLANVICKFKGEVIA-----KNGDVSRYRMDMEITNDGTIDPTE 318
Db      2266 ---DTLEDQL-----REVLPVAPSSSDIPFY-----STVTGGGLDTAE 2301

QY      319 DTFGPIGSPDFOGILFGVASQRKNEQNQPATRAHEAIINTGGDHLCPQISSSEIYLTSPN 378
Db      2302 -----LDADYWRNMRREPVEFEKATALLADGHD-----VFLE--- 2334

QY      379 ILRCTNPQPLQSGLRGTILIRSDNGHCHDMVGT----- 412
Db      2335 ----TSPHMLAISLQETI---SDAGASAAVLGTLRRGGGPRWLGVAVCRAYTHGV EID 2387

QY      413 ---SPTT---PTWPOWR-----CSRGSNCCSSG-----H 437
Db      2388 ABALFGDPRPVGLPTYPEQRERYWYSPVSRGDDPASLGDLDAADHPLGGGVGELPGSGDQ 2447

QY      438 RYPVPVVMNRVTWIV-----LSHKSGFSTSTRKLPQLNLRWPLIRFTN 480
Db      2448 MYTARIGTDAVPWLVDHALMGTVLLPGAVFTDLALWAGRQTGTGRIEELTLAAPLVL--- 2504

QY      481 PDTGRVLF-----EAR-----LHKQGITVAHTGDNP---IVMPP 512
Db      2505 PESGGVWLRNLNVGAPDTEARRFAVHARPEGAADWTLHAEGLLTAHEAADAADPASA VTPS 2564

QY      513 NGYFR-----FEAWNQFYS LAPVGTG 534
Db      2565 HGAEQLDGTGDFYERFTELGYSGYGPFRG 2592

RESULT 13
US-10-258-951-74
```

```

; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 66008
; LENGTH: 861
; TYPE: PRT
; ORGANISM: Neisseria meningitidis
; US-10-282-122A-66008

Query Match      3.7%; Score 108.5; DB 12; Length 861;
Best Local Similarity 19.3%; Pred. No. 0.86;
Matches 117; Conservative 77; Mismatches 238; Indels 173; Gaps 26;

QY      1 MKMASNDAASSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNNFVOAP 60
Db      46 VELLARHPVPCVDDAAKVKASP-----LAAVAEGSASP-----LISPEVATELLGTM 93

QY      61 AGEFTVSPRNSPGEILLDLLELGPDLNPYLAHLARMYNGHAGMEVOIVLAGNAFTAGKII 120
Db      94 LGGYNI---HALIELLDDDKLAPAAKGLKHTLLMFDSDHVOEK---AEKGNKYAQEVILQ 148

QY      121 FAAPGPPFVENSQITMCPHVIIIDVRQLEPFLPMPDIWNPFHYNQNDPKRLVA 180
Db      149 SWADAEMFASRAKVEKITVTVFK-VDGETNTDDLSAPADWSR-----PDIPLHALA 200

QY      181 MLXTP-----LRANNS-----GDDVFTVSCR-----VLTKPSPD 209
Db      201 MLKNPRDGIPTDKEGVEGPIKLEELKAKGHPVAVYGVGVGTGSSRSKATNSVIWHTGED 260

QY      210 FEFTFLVPPTVESKTKQFA-----LPILKISEMTNSRFPVVDVMYTARNENQVVOQP 261
Db      261 IPF---VP-----NKRFGVCLGGKIAPIFFNTQEDSGALPIEDVDVSALKMGDVVDILP 311

QY      262 QNGRVTLDELGLGTTPLLA-----VNICKFKGEVIAKNGDVSRYRMDMEITNT 309
Db      312 YEGKIVNGETVAEFELKSVLLDEVOAGGRINLIIRG-LTAXAREA---LKLPASTA 366

QY      310 DGTPIDPTEDTPG-----PIGSPDFQGLIFG-----VASQRKNEQNPA TR- 350
Db      367 FRLPQAPAESKAGFTLAQKMWGRACGLPEGQGVRRPGTYCEPRMTTVGSQDTGPMTRDEL 426

QY      351 AHEAIINTGGDHLCPQISSSEIYLTSPNILECTNPQLPQ-SGLRGITILIRSDNGHCH-- 407
Db      427 KDLACLGFSA DMVMSQSFCHTAAY---PKPVDVTKHKELPAFISTRGVSLRPGDVHSHW 483

QY      408 -----DMVGTSTPTPTWPOWRRCRSGNSCCSSGHRYPVPVVMNRVTWIVLSHKS GFS 460
Db      484 LNRLLLPDTVTGDSHT-----RFGISFPAGSLV-----AFA 519

QY      461 TSTRKLPQLNLRWPLIRFINPDTGRVLFPEARLH-----KQGITVAHT 503
Db      520 AATGWMP-LDM-----PESVLRVPSGKLQPGVTLRLDVNAIPLYAIKQGLITVAKA 569

QY      504 GDNFI 508
Db      570 GKQNI 574

RESULT 12
US-10-229-148B-4
; Sequence 4, Application US/10229148B
; Publication No. US20040091975A1
; GENERAL INFORMATION:
; APPLICANT: Meiji Seika Kaisha, Ltd.
```

Sequence 74, Application US/10258951  
Publication No. US20040033504A1  
GENERAL INFORMATION:  
APPLICANT: Agarwal, Pankaj  
APPLICANT: Murdock, Paul R.  
APPLICANT: Rizvi, Safia K.  
APPLICANT: Smith, Randall F.  
APPLICANT: Xiang, Zhaoying  
APPLICANT: Kabnick, Karen  
APPLICANT: Lai, Ying-Ta  
APPLICANT: Xie, Qing  
TITLE OF INVENTION: NOVEL COMPOUNDS  
FILE REFERENCE: GP50025  
CURRENT APPLICATION NUMBER: US/10/258,951  
CURRENT FILING DATE: 2002-10-28  
PRIOR APPLICATION NUMBER: PCT/US01/13360  
PRIOR FILING DATE: 2001-04-26  
PRIOR APPLICATION NUMBER: 60/199,963  
PRIOR FILING DATE: 2000-04-27  
PRIOR APPLICATION NUMBER: 60/203,336  
PRIOR FILING DATE: 2000-05-11  
PRIOR APPLICATION NUMBER: 60/207,087  
PRIOR FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: 60/207,546  
PRIOR FILING DATE: 2000-05-26  
NUMBER OF SEQ ID NOS: 78  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 74  
LENGTH: 607  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-258-951-74

Query Match 3.7%; Score 107; DB 12; Length 607;  
Best Local Similarity 19.4%; Pred. No. 0.7;  
Matches 116; Conservative 73; Mismatches 208; Indels 200; Gaps 29;  
QY 51 WIRNFFVQAPAGFTVSPRNSPGEILL-----DLELGPDLNPLYLAHARMYNGHAGMEV 105  
DB 46 WYWNQF-----FVIEYAGPEPVLIKLSHDVDRGERTKYL-----LTGEGAGTVFV 93  
QY 106 QIVLAGNAFTAGKIIFAAIPPGFPYENLSPSITMCPHVI--IDVRLQEP--FLLPMPD 160  
DB 94 IDEATGNHVTKSL-----DREEKAQVLLAQAADVADRASNRPLEPPSEFIKQVD 142  
QY 161 IWNN-----PFHYNQGNPKLRLVAMLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFT 213  
DB 143 INDNPIFPLGPHY---ATVPMSNVGTSVIQVTAHADDDPSYGSAKLVYTVLDGLPFF 199  
QY 214 FLVP-----PTVESKTKQFALPKISEMTNS-----RPFVPVDMYMTARNENQVVQ 260  
DB 200 SVDPTQGVVRTAIPNMDRETQBEFLVVIQAKDMGHHMGLSGSTTVTLSDVNDNPPKF 259  
QY 261 PQN-----GRVTLDEL-----LGTTPLLAVNICFKPG-EVIAKNGDV----- 297  
DB 260 POSLYQFSVETAGPGLVRLRAQDPDILGDNALMAYSILDGSEAFSISTDLOGRDL 319  
QY 298 -----RSYRMDMEITNDGTPTDP-----TEDTPGPIGSPD 328  
DB 320 LTVRKPLDFESQRSYFRVEATN---TLIDPAYLRGPPKDVASVRVAVQDAPEP---PA 373  
QY 329 FQGLFGVASQORNEQNPAHAEAIINTGGDHLCPQISSEIYLTSPNILRCTNPQPL 388  
DB 374 FTQAAVHLTVPENK---AFGT-----LVGQISAADL-----DS 403  
QY 389 POSGLRGTLIRSDNGHC-----HDMVCTS-----PTPTTW-- 419  
DB 404 PASPIRYSILPHSDPERCFSIQPEEGTIHTAAPLDREARAWNLTVLATELVYPTPAYAS 463  
QY 420 --PQWRRCRCSGNC--CSSGH-RYPVPVVMNVTVM-----IVLSHSGFSTSTRKL---PQ 468  
DB 464 GAPPPECLHTAYENCPCICGYLNVSKAYMNVMMWMLVFAEHKGGGGRGQAVDQOK 523

QY 469 LNLRWPLIRFIN-----PDTGRVLFEARL-----HKQGITVVAHTGDNPIV 509  
DB 524 QSTRWELSONCNYLPKSRGVHPGTSMRAQASQOQSRGTHRNCTRIACHTRVNPIL 580  
RESULT 14  
US-10-296-115-1211  
Sequence 1211, Application US/10296115  
Publication No. US20040053248A1  
GENERAL INFORMATION:  
APPLICANT: Hyseq Inc  
TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides  
FILE REFERENCE: 784PCT  
CURRENT APPLICATION NUMBER: US/10/296,115  
CURRENT FILING DATE: 2002-11-18  
PRIOR APPLICATION NUMBER: US09/488,725  
PRIOR FILING DATE: 2000-01-21  
PRIOR APPLICATION NUMBER: US09/552,317  
PRIOR FILING DATE: 2000-04-25  
NUMBER OF SEQ ID NOS: 1478  
SEQ ID NO 1211  
LENGTH: 767  
TYPE: PRT  
ORGANISM: Homo sapiens  
US-10-296-115-1211

Query Match 3.7%; Score 106.5; DB 12; Length 767;  
Best Local Similarity 20.0%; Pred. No. 1.1;  
Matches 112; Conservative 65; Mismatches 200; Indels 183; Gaps 30;  
QY 51 WIRNFFVQAPAG-EFTVSPRNSPGEILLDLLELGPDLNPLY--LAHARMYNGHAGMEVQI 107  
DB 204 WKKDNEVLTNADMENFVHVHAQDGEVM-----EYTTILHLRVQTFGHEG--RYQC 251  
QY 108 VLA---GNAFTAGKIIFAAIPGFPYENLSPSITM-----C-----BHVLIID- 147  
DB 252 VITNHFGSTYSHKARLTNVNLPF---TKTPHDITIRTTMARLSAATGHNPOJAWOK 308  
QY 148 -----VRQLEPFLPMPDIWNNPFHYNQGNPKLRLVAMLYTPLRANNSGDDVFTVS 199  
DB 309 DGGTDFPAARERRMHVMPDDV---FFITDVKIDD-----AGVYCTAQSAGSISANAT 360  
QY 200 CRVLTKPS---PDFEFTFLVPTVESKTKQFALPKISEMTNSRFPVDMYMTARNEN 256  
DB 361 LTVLTPSLVPLDRVVSVGETVALQCKATGNPPRITWPKGDR-PLSLTERHHLTPDN 419  
QY 257 QVVOQPN-----GRVTLD-CELLGT-----TPLLAVNICFKGEVIAKNG- 295  
DB 420 QLLVVQNVVAEDAGRYTCMSENTLGTERRAHSQLSVLPAAGCKKOGTTV---GIFTIIVVS 476  
QY 296 -----DVSRYRMDMEITNDGTPTDPIDPTEDTPGPIGSPDFQGLFGVASORN 341  
DB 477 SILVTSLVVWCIIYQTRKKSEYVNTDETVPVP--DVPSYLS--QGTLSDRQETVV 531  
QY 342 KNEQNPAHAEAIINTGGDHLCPQISSEIYLTSPNI--LRCTNPQ----- 386  
DB 532 RTEGGPQANGH---IESNG--VCPRDAS---HFPEPDTHSVACRQPKLCAGSAVHKPKWK 583  
QY 387 -----PLQSGLRGTILIRSDNGHCHDMV-----GTSPTTPTWPQ 421  
DB 584 AMEKAEGTGPCHKMEHGRVVCSD---CNTEVDVCSYRGOAFHPQVPSRDSQAQPSAPNGE 640  
QY 422 -----QWRRCRSGNCCSSGHHRYFPVPMNVMTVILSHKSGFSTSTRKL 469  
DB 641 PGGSQDEHSPHHQCSRTAAGSCPCCGSLYP-----SNHDBMLTAVKCKPMA 687  
QY 470 NL-----RWPLIRFINPDT 483  
DB 688 SLDGKGDSSWTLARLYHPDS 707

RESULT 15  
US-10-282-122A-65330

Db	312	YEGKVNKQETVAEFELKSQVLLDEVOAGGRINLIIRG-LTKAREA----	YEGKVNKQETVAEFELKSQVLLDEVOAGGRINLIIRG-LTKAREA----	YEGKVNKQETVAEFELKSQVLLDEVOAGGRINLIIRG-LTKAREA----	YEGKVNKQETVAEFELKSQVLLDEVOAGGRINLIIRG-LTKAREA----
QY	310	DGTRIDPDPEDTPG-----PIGSPDPQILFG-----	DGTRIDPDPEDTPG-----PIGSPDPQILFG-----	DGTRIDPDPEDTPG-----PIGSPDPQILFG-----	DGTRIDPDPEDTPG-----PIGSPDPQILFG-----
Db	367	FLPQAPAEKVGFTLQAKVGRACGLPEQGVPGTYCEPRMTVGSODTTPWTRDEL	FLPQAPAEKVGFTLQAKVGRACGLPEQGVPGTYCEPRMTVGSODTTPWTRDEL	FLPQAPAEKVGFTLQAKVGRACGLPEQGVPGTYCEPRMTVGSODTTPWTRDEL	FLPQAPAEKVGFTLQAKVGRACGLPEQGVPGTYCEPRMTVGSODTTPWTRDEL
QY	351	AHEAIINTGGDHLCPQISSSEIYLTSPNILRCTNPQPLQ-SGLRGTILIRSDNGHCH--	AHEAIINTGGDHLCPQISSSEIYLTSPNILRCTNPQPLQ-SGLRGTILIRSDNGHCH--	AHEAIINTGGDHLCPQISSSEIYLTSPNILRCTNPQPLQ-SGLRGTILIRSDNGHCH--	AHEAIINTGGDHLCPQISSSEIYLTSPNILRCTNPQPLQ-SGLRGTILIRSDNGHCH--
Db	427	KDLACLGFSDAMVQMSFCHTAAY---PKPVDVTKHKLPAFISTRGGVSLRCDGVVHSW	KDLACLGFSDAMVQMSFCHTAAY---PKPVDVTKHKLPAFISTRGGVSLRCDGVVHSW	KDLACLGFSDAMVQMSFCHTAAY---PKPVDVTKHKLPAFISTRGGVSLRCDGVVHSW	KDLACLGFSDAMVQMSFCHTAAY---PKPVDVTKHKLPAFISTRGGVSLRCDGVVHSW
QY	408	-----DMVGTSTPTPTPWQOWRRCSRGNSCCSSGHYPVPVVMNRVTWIVLSHKSGFS	-----DMVGTSTPTPTPWQOWRRCSRGNSCCSSGHYPVPVVMNRVTWIVLSHKSGFS	-----DMVGTSTPTPTPWQOWRRCSRGNSCCSSGHYPVPVVMNRVTWIVLSHKSGFS	-----DMVGTSTPTPTPWQOWRRCSRGNSCCSSGHYPVPVVMNRVTWIVLSHKSGFS
Db	484	LNRLLLPTVGTGGDSHT-----RFPGISFPAGSLV-----	LNRLLLPTVGTGGDSHT-----RFPGISFPAGSLV-----	LNRLLLPTVGTGGDSHT-----RFPGISFPAGSLV-----	LNRLLLPTVGTGGDSHT-----RFPGISFPAGSLV-----
QY	461	TSTRKLPOLNLRWPLRFINPDTGRVLFEARLH-----	TSTRKLPOLNLRWPLRFINPDTGRVLFEARLH-----	TSTRKLPOLNLRWPLRFINPDTGRVLFEARLH-----	TSTRKLPOLNLRWPLRFINPDTGRVLFEARLH-----
Db	520	AATGVIP-LDM-----PESVLVRFSGKLOPGVTLRDLVNAIPLYAIKQGLLTVAKA	AATGVIP-LDM-----PESVLVRFSGKLOPGVTLRDLVNAIPLYAIKQGLLTVAKA	AATGVIP-LDM-----PESVLVRFSGKLOPGVTLRDLVNAIPLYAIKQGLLTVAKA	AATGVIP-LDM-----PESVLVRFSGKLOPGVTLRDLVNAIPLYAIKQGLLTVAKA
QY	504	GDNPI 508	GDNPI 508	GDNPI 508	GDNPI 508
Db	570	GKKN1 574	GKKN1 574	GKKN1 574	GKKN1 574

Search completed: June 1, 2004, 14:04:48  
Job time : 36.5841 secs

```

/ Sequence 65330, Application US/10282122A
/ Publication No. US20040029129A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, Liangsu
/ APPLICANT: Zamudio, Carlos
/ APPLICANT: Malone, Cheryl
/ APPLICANT: Haselbeck, Robert
/ APPLICANT: Ohlsen, Kari
/ APPLICANT: Zyskind, Judith
/ APPLICANT: Wall, Daniel
/ APPLICANT: Trawick, John
/ APPLICANT: Carr, Grant
/ APPLICANT: Yamamoto, Robert
/ APPLICANT: Forsyth, R.
/ APPLICANT: Xu, H.
/ TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
/ FILE REFERENCE: ELITRA.034A
/ CURRENT APPLICATION NUMBER: US/10/282,122A
/ CURRENT FILING DATE: 2003-02-20
/ PRIOR APPLICATION NUMBER: 60/191,078
/ PRIOR FILING DATE: 2000-03-21
/ PRIOR APPLICATION NUMBER: 60/206,848
/ PRIOR FILING DATE: 2000-05-23
/ PRIOR APPLICATION NUMBER: 60/207,727
/ PRIOR FILING DATE: 2000-05-26
/ PRIOR APPLICATION NUMBER: 60/230,335
/ PRIOR FILING DATE: 2000-09-06
/ PRIOR APPLICATION NUMBER: 60/230,347
/ PRIOR FILING DATE: 2000-09-09
/ PRIOR APPLICATION NUMBER: 60/242,578
/ PRIOR FILING DATE: 2000-10-23
/ PRIOR APPLICATION NUMBER: 60/253,625
/ PRIOR FILING DATE: 2000-11-27
/ PRIOR APPLICATION NUMBER: 60/257,931
/ PRIOR FILING DATE: 2000-12-22
/ PRIOR APPLICATION NUMBER: 60/267,636
/ PRIOR FILING DATE: 2001-02-09
/ PRIOR APPLICATION NUMBER: 60/269,308
/ PRIOR FILING DATE: 2001-02-16
/ Remaining Prior Application data removed - See File Wrapper or PALM.
/ NUMBER OF SEQ ID NOS: 78614
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 65330
/ LENGTH: 861
/ TYPE: PRT
/ ORGANISM: Neisseria gonorrhoeae
/ US-10-282-122A-65330

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:36:02 ; Search time 10.4388 Seconds  
(without alignments)  
4985.230 Million cell updates/sec

Title: US-09-926-799-11  
Perfect score: 2894  
Sequence: 1 MKMASNDRAAPSDGAAGLV.....VNQFYSLAPVGTGKRRRVQ 541

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78:\*  
1: Piri:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1620.5	56.0	542	2 S60616	capsid protein - h
2	1562	54.0	539	2 S40111	capsid protein - h
3	1004.5	34.7	530	2 B37471	capsid protein - N
4	966	33.4	546	2 B37491	major capsid prote
5	317	11.0	2344	1 RRRWRH	genome polyprotein
6	304	10.5	2344	2 S55399	genome polyprotein
7	304	10.5	2344	2 S64740	genome polyprotein
8	279.5	9.7	576	2 A53982	capsid protein - E
9	271.5	9.4	703	1 C48562	coat protein - San
10	262.5	9.1	671	1 VCVWF9	coat protein - fel
11	262	9.1	702	1 A48562	coat protein - San
12	260	9.0	668	1 VCVWFF	coat protein - fel
13	254	8.8	668	1 JCWFFF	capsid protein - f
14	248.5	8.6	668	2 JQ2354	capsid protein - f
15	248.5	8.6	668	2 JQ2356	capsid protein - f
16	126.5	4.4	2194	1 GNNYB7	genome polyprotein
17	126	4.4	3085	2 T00327	polyprotein - infe
18	118.5	4.1	2333	1 GNNY2F	genome polyprotein
19	117.5	4.1	2205	1 GNNY2W	genome polyprotein
20	117.5	4.1	2303	2 S13554	genome polyprotein
21	117	4.0	2175	1 GNNYBE	genome polyprotein
22	115.5	4.0	2207	1 GNNY1P	genome polyprotein
23	115.5	4.0	2209	1 GNNY2P	genome polyprotein
24	114	3.9	630	2 T03582	probable replicati
25	113.5	3.9	2206	1 GNNY21	genome polyprotein
26	112.5	3.9	2303	1 GNNYTP	genome polyprotein
27	112	3.9	2185	1 GNNYSH	genome polyprotein
28	112	3.9	2301	1 GNNYTN	genome polyprotein
29	112	3.9	2332	1 GNNY4F	genome polyprotein

30	111	3.8	2183	1 GNNYB4	genome polyprotein
31	110	3.8	2209	1 GNNY3P	genome polyprotein
32	109	3.8	929	2 A44048	genome polyprotein
33	109	3.8	2185	1 GNNYSV	genome polyprotein
34	109	3.8	2290	1 GNNYE	genome polyprotein
35	108.5	3.7	861	2 A81801	aconitate hydratase
36	108.5	3.7	2207	1 GNNYSP	genome polyprotein
37	108	3.7	2206	2 S03822	genome polyprotein
38	108	3.7	13055	2 T16580	hypothetical prote
39	106.5	3.7	940	2 D89723	protein F39D8.1b (
40	106.5	3.7	945	2 T21998	hypothetical prote
41	106.5	3.7	2481	2 A43908	fibronectin - Afri
42	106	3.7	2201	1 GNNYA9	genome polyprotein
43	106	3.7	2206	1 GNNY4P	genome polyprotein
44	106	3.7	2207	2 S09553	genome polyprotein
45	104.5	3.6	2303	1 GNNYTM	genome polyprotein

ALIGNMENTS

RESULT 1

S60616

capsid protein - human calicivirus (strain Melksham)

C:Species: human calicivirus

A:Variety: strain Melksham

C:Date: 23-May-1997 #sequence\_revision 23-May-1997 #text\_change 28-Jul-2000

C:Accession: S60616

R:Green, S.M.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.

Virus Res. 37, 271-283, 1995

A:Title: Capsid diversity in small round-structured viruses: molecular characterization

A:Reference number: S60615; MUID:96136658; PMID:8533462

A:Accession: S60616

A:Molecule type: genomic RNA

A:Residues: 1-542 <GRE>

A:Cross-references: EMBL:X81879; NID:G976077; PIDN:CAA57462.1; PID:G976079

A:Experimental source: strain Melksham

A:Note: it is uncertain whether Met-1 or Met-3 is the initiator

C:Superfamily: human calicivirus capsid protein

C:Keywords: capsid protein; coat protein

Query Match 56.0%; Score 1620.5; DB 2; Length 542;  
Best Local Similarity 58.1%; Pred. No. 28-116;  
Matches 324; Conservative 68; Mismatches 133; Indels 33; Gaps 7;

Qy 1 MKMASNDRAAPSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNNFVQAP 60

Db 1 MKMASNDRAAPSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNNFVQAP 60

Qy 61 AGEFTVSPRNSPGEILLDLELGPDLNPLYLAHLMRYNGHAGMEVQIVLAGNAFTAGKII 120

Db 61 NGEFTVSPRNPAPGVEVLLNLELGPDLNPLYLAHLMRYNGHAGMEVQIVLAGNAFTAGKII 120

Qy 121 FAATPGPFYENLSPQITMCPHVIIDVROLEPPLPMPDIWNFFHYNQNDPKRLVA 180

Db 121 FAAPVPHFPVENLSPQITMCPHVIIDVROLEPPLPMPDIWNFFHYNQNDPKRLVA 180

Qy 181 MLYTPLRANSGDDVFTVSCRVLTGKPSDFEFTFLVPPVTESKTKQFALPILKISEMNTS 240

Db 181 MLYTPLRANSGDDVFTVSCRVLTGKPSDFEFTFLVPPVTESKTKQFALPILKISEMNTS 240

Qy 241 RFPVVDVVMYTARNENQVVPQNGRVTLTDLGELLGTTPLLVANNICKFKGEVIAKNGDVRSY 300

Db 241 RFPVVDVVMYTARNENQVVPQNGRVTLTDLGELLGTTPLLVANNICKFKGEVIAKNGDVRSY 300

Qy 301 RMDMEITNDGTPDPTDTPGPGSPDFQIGLFGVASQRNK-----NQONPATRAHAII 356

Db 300 LNNVITITNLGSPDFPDSIDIAPLGVDPFGVRFGVVISQDKQNAAGHSEPPANRGHDVAV 359

Qy 357 NTGGDHLCPQISSSEIVLTSNILLRCTNPQPLPQSGLRGTLIRSDNGHCHDMV----- 410

Db 360 PTYTAQYTPKLGQIQIGTQWTDLTLVNQVFKVTPVGL-----NDTEHFNQVMPRVAG 412

```
QY 411 -----GTSPT-TPTWPOWRRCRSGNCCSSGHRYPFVPMVNRVTWVLSHKGSFSTST 463
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 413 ALNLTNLTAPSVAPVPPGRLFFRSHLPKGGYGNPAIDCLLPQEWQHFYQEA----- 467
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 464 RKLQNLRLWPLIRFNDPTGVLFEARLHKOGFTTVAHTGDNPIVMPGNGYFRPEAWVN 523
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 468 --APSMS-EVALVYINPTDGRALFEAKUHRAGFMVSSNTSAPVVVPANGVFRDSDWN 524
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 524 QFYSLAPVGTGKRRRVQ 541
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 525 QFYSLAPMGNGRRRVQ 542
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 2
S40111
capsid protein - human calicivirus (strain Bristol isolate B493)
C:Species: human calicivirus
A:Variety: strain Bristol isolate B493
C:Date: 25-Dec-1994 #sequence_revision 27-Feb-1997 #text_change 28-Jul-2000
C:Accession: S40111
R:Green, S.M.; Dingle, K.E.; Lambden, P.R.; Caul, E.O.; Ashley, C.R.; Clarke, I.N.
submitted to the EMBL Data Library, December 1993
A:Description: Human enteric caliciviridae: a new prevalent SRSV group defined by RNA-de
A:Reference number: S40111
A:Accession: S40111
A:Molecule type: genomic RNA
A:Residues: 1-539 <GRE>
A:Cross-references: EMBL:X76716; NID:G436410; PIDN:CAA54134.1; PID:G436411
A:Experimental source: human enteric calicivirus strain Bristol isolate B493
C:Superfamily: human calicivirus capsid protein
C:Keywords: capsid protein; coat protein

Query Match 54.0%; Score 1562; DB 2; Length 539;
Best Local Similarity 54.3%; Pred. No. 6e-112;
Matches 305; Conservative 76; Mismatches 133; Indels 48; Gaps 8;

QY 1 MKMASNDAASSDGAAGLVEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNRFVQAP 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MKVASNDANPSGSAANLVEVNEVNALEPVVGAIAAPVACQNVIDPWIRNRFVQAP 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 61 AGEFTVSPRNSPGEILDLDELGPDLPNLYLAHARMYNGHAGGMEVQIVLAGNAFTAGKII 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 GGEFTVSPRAPGEILLASAPLGPDLPNLYLSRMVNGYAGGVEVQIVLAGNAFTAGKVI 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 121 FAIIPGFPVENLSPQITMCQPHVIDVQLEPFLPMPDIWNNFHYNQNDPKRLVA 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 FAAPVPPNFTGLSPSQVTFPHIIVDVRQLEPVLIELPDVRNFFHYNQANDSTKLIA 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSDFEFTFLVPTVESKTKQFALPILKISEMTNS 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 MLYTPLRANAGDDVFTVSCRVLTRPSDFDFLFLVPTVESKTKPTVPLVTEEMSNS 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 241 RFPVPVDMVTARNENOVQPNQGRVTLDGELGTTTPLLAVNICKFKEGVIANKGDVRSY 300
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 RFPPILEKLYTGPSSAFVQPNQGRCTTDGVLGTTQLSAVNICNFRGVDTHIAG---SH 297
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 301 RMDMEITNDGTPIDPTEDTPGIPGDFQGLFQGVASQRNKEQNQNPATRAHEALINTGG 360
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 DYTWNLASQNSNYSNDEETEEIPAPLGTDFGVKIQGLLTQTR--ADGSTRAHKATVSTGS 355
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 361 DHLCPQISSSIYLTSPNLRCTNPQPLPQSGLRGTILIRSDNGHCHDMVGTSTPTTWP 420
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 356 VHFTPKLGVSQFTDNDFOAGQNTKFTPVG-----VIQGDHQN-----P 399
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 421 QQWRRCRSGNCCSGHRYP--VPVPMNRVTWVLSHKGSFSTSTRKLQNL-----R 472
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 400 QQWLLPNYSGRTGHNVLAPAVPTFGEQLLFFRSTMPGCS-----GYPNMNLCLLPQE 455
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 473 W-----PLIRINPDTRGVLFEARLHKOGFTTVAHTGDNPIVMPGNGYFR 517
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 456 WVLHFYQEAAPQASDVALLRVNDPTGRVLFECKLHKSGYITVAHTGIDVLVLPNGYFR 515
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 518 FEAWNQFYSLAPVGTGKRRR 539
```

```
Db 516 FDSWVNFYTLAPMGNGTGRRR 537
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 3
B37471
capsid protein - Norwalk virus
C:Species: Norwalk virus
C:Date: 24-Feb-1994 #sequence_revision 25-Apr-1997 #text_change 28-Jul-2000
C:Accession: B37471
R:Jiang, X.; Wang, M.; Wang, K.; Estes, M.K.
Virology 195, 51-61, 1993
A:Title: Sequence and Genomic organization of Norwalk virus.
A:Reference number: A37471; MUID:93303939; PMID:8391187
A:Accession: B37471
A:Status: preliminary; not compared with conceptual translation
A:Molecule type: genomic RNA
A:Residues: 1-530 <JTA>
A:Cross-references: GB:M87661; NID:G1061311; PIDN:AAB50466.1; PID:G1061313
A:Note: sequence extracted from NCBI backbone (NCBIP:134157)
C:Superfamily: human calicivirus capsid protein

Query Match 34.7%; Score 1004.5; DB 2; Length 530;
Best Local Similarity 40.8%; Pred. No. 3.6e-69;
Matches 234; Conservative 73; Mismatches 180; Indels 87; Gaps 14;

QY 1 MKMASNDAASSDGAAG--LVPEIN-NEVMPLEPVAGASLATPVVQQNIIDPWIRNRF 56
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1 MMMAKSDATSSVDGASGAGQLVPEVNASDPLAMPFVAGSSTAVATAGQVNPIDWIINN 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 57 VQAPAGEFTVSPRNSPGEILDLDELGPDLPNLYLAHARMYNGHAGGMEVQIVLAGNAFTA 116
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 61 VQAPQGEFTSPNNTPGDVLFDLSLGPHLNFLLHLSQMYNGVGNMVRIMLAGNAFTA 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 117 GKIIIPAAIPPGFPVENLSPQITMCQPHVIDVQLEPFLPMPDIWNNFHYNQNDPKL 176
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 121 GKIIIVSCIPPGFGSHNLTIAQATLFFPHVIAVDTLDPIEVLVEDVRNVLFHNNDRNQTM 180
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 177 RLVMALYTPLRANSGDDVFTVSCRVLTKPSDFEFTFLVPTVESKTKQFALPILKISE 236
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 181 RLVCMLYTPLRGTGGTGSFVVAGRVMTCPSPDFNFLFLVPTVEQKTRPFTLPNPLSS 240
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 237 MTNSRFPVPMVTARNENOVQPNQGRVTLDGELGTTTPLLAVNICKFKEGVIANKGD 296
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 241 LNSRAPLPISSIGISPDNVQSPQNGRCTLGDLGVTGTPVLSHVAKIRG---TSNGT 297
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 297 VRSYRMDMEITNDGTPIDPTEDTPGIPGDFQF-----ILFGVASQRNKEQNQNPAT 349
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 298 V-----INLTEDGTPHPFEG-PAPIGFDLGGCDWHINMTQFGHSSQIQYD----- 344
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 350 RAHEALINTGGDHLCPQISSSIYLTSPNLRCTNPQPLPQSGLRGTILIRSDNGHCHDM 409
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 345 -----VDITPTDTPVPHLGS-----IQANGICSGNY 369
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 410 VG-----TSPTPTTPWQ--QWRRCRSGNCCSGHRYP--VPVPMNRVTWVLSHKGSFS 460
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 370 VGVLSWISPPSPSGSQVDLMKIPNYGSSITATHLAPSVYPPGFEVLVFFMCKMPGPG 429
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 461 TSTRK--LPQLNL-----RWPLIRINPDTRGVLFEARLHKOGFTTVAHTG--- 504
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 430 AYNLPCLLQFEYISHLASQAQTVGEAALLHYVDPTGRNLGFEKRAYPDGFLTCVPNGAS 489
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 505 DNPVMPNGYFRFEAWNQFYSLAPVGTGKRR 538
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 490 SGPQQLPINGVFVFSWVSFRYQLKPVGTASSAR 523
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 4
B37491
major capsid protein [similarity] - Southampton virus
N:Alternate names: orf2 protein
C:Species: Southampton virus
C:Date: 22-Sep-1993 #sequence_revision 25-Apr-1997 #text_change 18-Aug-2000
```



```
QY 89 LAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPGCFPYENLSPS-QITMCPHVLIID 147
Db 1868 TAVLSQMYAGWAGMQFRFIVAGSGVFGRLVAIVPPGI---EIGPGLVRFQPHVVVID 1924
QY 148 VRQLEPFLPMPDIWNFFHYNQGNPKL--RLVAMLYTPLRANNSGDDVFTVSCRVLTK 205
Db 1925 ARSLPEVTITMPDLRPNMYH--PTGDPGLVPLVLVSVINNL-INPFGGSTSAIQVTVETR 1981
QY 206 PSPDFEFTFLVPTVESKTKQFALPILKISEMTNSRFPVPVDMVTARNENQVV--QPON 263
Db 1982 PSEDFEFVMIRAPS--SKT-----VDSISPAGLLTTPVLTVGVGNDNRWNGQIVGLQFVP 2033
QY 264 GRVT-----LDGELLG-TTPLLAVNICFKPGEVIKNGDVRSYRMDMEITN----- 308
Db 2034 GGFSTCNRHWNLSGTYGWSSPFA-----DIDHRRGSA-SY-PGSNATNVLOFWYA 2083
QY 309 TDGTPIDPTEDTPGPIGSPD-----FOG-----ILFGVASQRNKNQONPAT-RAHEA 354
Db 2084 NAGSAVDNFIQVAPDGFDMGSFVPGGIPAGCHVGFAGIWNNSGAPNVTTVOAYEL 2143
QY 355 IINTGG-DHLCQFQISSSEIYLTSPNLRCTNPQPLQSGL-----RGTLIRSDNGHCH-- 407
Db 2144 GFATGAPGNLQPTNTSGQTVAKSIYAVVTGTAQNPAGLFVMASGVISTPNANAITYP 2203
QY 408 --DMVGTSTPTT 417
Db 2204 QPDRIVTTTGT 2215

RESULT 7
S64740
genome polyprotein - rabbit hemorrhagic disease virus (isolate AST/89)
N;Contains: VP60 protein
A;Variety: isolate AST/89
C;Date: 12-Jul-1996 #sequence revision 26-Jul-1996 #text_change 20-Jun-2000
C;Accession: S64740; S64944; S49018; S65012
R;Casais, R.; Martin-Alonso, J.; Boga, J.; Parra, F.
submitted to the EMBL Data Library, May 1995
A;Description: Genomic organization of rabbit hemorrhagic disease virus determined by di
A;Reference number: S64740
A;Accession: S64740
A;Molecule type: genomic RNA
A;Residues: 1-2344 <CAS>
A;Cross-references: EMBL:Z49271; NID:g1182032; PIDN:CAA89265.1; PID:g1150552
R;Boga, J.; Casais, R.; Marin, M.; Martin-Alonso, J.; Carmenes, R.; Prieto, M.; Parra, F.
submitted to the EMBL Data Library, July 1993
A;Description: Molecular cloning, sequence and expression of the capsid protein gene fr
A;Reference number: S46944
A;Accession: S46944
A;Molecule type: genomic RNA
A;Residues: 1650-2344 <BOG>
A;Cross-references: EMBL:Z24757; NID:g515622; PIDN:CAA80881.1; PID:g515623
R;Parra, F.; Boga, J.A.; Marin, M.S.; Casais, R.
Virus Res. 27, 219-228, 1993
A;Title: The amino terminal sequence of VP60 from rabbit hemorrhagic disease virus supp
A;Reference number: S49018; MUID:93255896; PMID:8488721
A;Accession: S49018
A;Molecule type: genomic RNA
A;Residues: 1650-1796 <PAR>
A;Cross-references: EMBL:Z24757
A;Experimental source: EMBL:Z24757
A;Accession: S65012
A;Molecule type: protein
C;Residues: 1767-1779;1875-1877,'X',1879-1881;1936-1938,'X',1940-1941 <PAW>
C;Superfamily: rabbit calicivirus RNA-directed RNA polymerase
C;Keywords: blocked amino end; polyprotein

Query Match 10.5%; Score 304; DB 2; Length 2344;
Best Local Similarity 29.6%; Pred. No. 1.7e-14;
Matches 128; Conservative 55; Mismatches 175; Indels 74; Gaps 24;
```

```
QY 33 AGASLATPVVG---QQNIIDPWIRNMFVQAPAGEFTVSPRNSPGEILLDLGLDPLNYP 88
Db 1811 SSASIIATAGTGGPQQVDDQETW-RTFYNDV--FTWSVADAPGSLTYTVQHSPPNNPF 1867
QY 89 LAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPGCFPYENLSPS-QITMCPHVLIID 147
Db 1868 TAVLSQMYAGWAGMQFRFIVAGSGVFGRLVAIVPPGI---EIGPGLVRFQPHVVVID 1924
QY 148 VRQLEPFLPMPDIWNFFHYNQGNPKL--RLVAMLYTPLRANNSGDDVFTVSCRVLTK 205
Db 1925 ARSLPEVTITMPDLRPNMYH--PTGDPGLVPLVLVSVINNL-INPFGGSTSAIQVTVETR 1981
QY 206 PSPDFEFTFLVPTVESKTKQFALPILKISEMTNSRFPVPVDMVTARNENQVV--QPON 263
Db 1982 PSEDFEFVMIRAPS--SKT-----VDSISPAGLLTTPVLTVGVGNDNRWNGQIVGLQFVP 2033
QY 264 GRVT-----LDGELLG-TTPLLAVNICFKPGEVIKNGDVRSYRMDMEITN----- 308
Db 2034 GGFSTCNRHWNLSGTYGWSSP-----RF-GDIGHRRGSA-SYPGN-NATNVLOFWYA 2083
QY 309 TDGTPIDPTEDTPGPIGSPD-----FOG-----ILFGVASQRNKNQONPAT-RAHEA 354
Db 2084 NAGSAVDNFIQVAPDGFDMGSFVPGGIPAGCHVGFAGIWNNSGAPNVTTVOAYEL 2143
QY 355 IINTGG-DHLCQFQISSSEIYLTSPNLRCTNPQPLQSGL-----RGTLIRSDNGHCH-- 407
Db 2144 GFATGAPGNLQPTNTSGQTVAKSIYAVVTGTAQNPAGLFVMASGVISTPNANAITYP 2203
QY 408 --DMVGTSTPTT 417
Db 2204 QPDRIVTTTGT 2215
```

## RESULT 8

```
A53982
capsid protein - European brown hare syndrome virus
C;Species: European brown hare syndrome virus
C;Date: 16-Jul-1999 #sequence_revision 16-Jul-1999 #text_change 28-Jul-2000
C;Accession: A53982
R;Wirblich, C.; Meyers, G.; Ohllinger, V.F.; Capucci, L.; Eskens, U.; Haas, B.; Thiel, H.
J. Virol. 68, S164-S173, 1994
A;Title: European brown hare syndrome virus: relationship to rabbit hemorrhagic disease
A;Reference number: A53982; MUID:94309183; PMID:7518531
A;Accession: A53982
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-576 <WIR>
A;Cross-references: GB:U09199
C;Superfamily: human calicivirus capsid protein
```

Query Match 9.7%; Score 279.5; DB 2; Length 576;  
Best Local Similarity 25.6%; Pred. No. 1.7e-13;  
Matches 121; Conservative 60; Mismatches 170; Indels 121; Gaps 23;

```
QY 17 GLVPEI--NNEYMPLFVAGASLATPVVG-----QQNIIDPWIRNMFVQAPAGEFTVSPRN 70
Db 25 GMDPGVAGTDDVTADNVA-ASVATAGIGGPPQASQPSQSRVNFYFN---DVFTWSVTD 80
QY 71 SGEILLDLGLDPLNLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPGFPY 130
Db 81 AFGSILYSVQHSFPQNNPFTQVLSQMYAGWAGMQFRFIVAGSGIFGGRVLCALIPPGI-- 138
QY 131 ENLSPS-QITMCPHVLIIDVRQLEPFLPMPDIWNFFHYNQGNPKL--RLVAMLYTPLR 187
Db 139 -QIQPGLVRFQPHVVIDARSLEPVTITMPDLRPNMYH--PTGDPGLVPLTVSVVNNI- 194
QY 188 ANNSGDDVTVCRCVLTKSPDFEFTFLVPP---TVESKTKQFAL--PILKISEMTNSRF 242
Db 195 INPFGGTSAIQVTVETRESEDFEFLIRAPSKTVDSVNPSSLTTPVL-TCAGSDNRW 253
QY 243 FVPVDVMTARNENQVVQP-----QNGRVITLDGELLGTTPLAVNICFKPGEVIKNG 295
```



Db 254 GAPI-----VGLQPVGGFSTSNRHNWNGSTYGSWSPRFPDIDHPSGNVSPYTG 303  
QY 296 DVRSYRMDMEITWT-----DGTPTDPTEDTPGPIGSPDFOGILF-----GVA 337  
Db 304 SA-----TNTIETWYANAGATTATNPISNIAPDGPFDGAIPTSGTTIPTGAWVGFG 354  
QY 338 SQNKNEQNP--ATPAHE-----AIINTGGDHLCP-----QISSSE 371  
Db 355 QWNASNGTFYVGTVOAYELGFANGAPSSIRPVTITTTGAQLVAKSYGYVAIAQNOSSAGI 414  
QY 372 IYL-----TSPNLRCT-NPQOPQSGGLRGTILRSDNGHCHDMVGTSTPTP 417  
Db 415 IFLSKMVGTFPGVAATYTPQ-----SAIVTTGTP 446

RESULT 9  
C48562  
coat protein - San Miguel sea lion virus (serotype 4)  
N:Alternate names: capsid protein  
C:Species: San Miguel sea lion virus  
C>Date: 17-Feb-1994 #sequence\_revision 17-Feb-1994 #text\_change 23-Jul-1999  
C:Accession: C48562  
R:Neill, J.D.  
Virus Res. 24, 211-222, 1992  
A:Title: Nucleotide sequence of the capsid protein gene of two serotypes of San Miguel sea lion virus.  
A:Reference number: A48562; MUID:92410750; PMID:1529644  
A:Accession: C48562  
A:Molecule type: genomic RNA  
A:Residues: 1-703 <NEI>  
A:Cross-references: GB:M87482; NID:g334886; PIDN:AAAL6220.1; PID:g334888  
A:Note: sequence extracted from NCBI backbone (NCBIP:113567)  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:89,208,329,463,482/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.4%; Score 271.5; DB 1; Length 703;  
Best Local Similarity 27.0%; Pred. No. 9.5e-13;  
Matches 106; Conservative 57; Mismatches 160; Indels 69; Gaps 18;

QY 8 AAPSSDGAAGLPEINNEWPLE---PVAGASLATPVVGGQNIIDPWIRNPFVQAPAGE 63  
Db 151 AESDGGCSAEIVTEEGTVVQQQPAPAPTALATLATASTG-KSVEQEWMTFFSYHTSINW 209  
QY 64 FTVSPRNSPGEILLDLELGLDLPYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAA 123  
Db 210 STV---ESQKILYQALNPSINPYLDHIAKLYSTWSGGIDVRFVSGSGVFGKLAALL 266  
QY 124 IPPGF-PYENLSPSQITMCPHVIIDVROLEPFLMPDIWNNFFHYNQGNPKLRVLVAML 182  
Db 267 VPPGVEPIESVSMLOQ---PHVLFDAQTEPVIFITPIDIRKTLFHSMDTD-TTKLVINP 322  
QY 183 YTPLRANSGDDVFTVSCVLTKPSDFEFTLVLPPTVESKTKQFALPILKISEMNSRF 242  
Db 323 Y-----ENGVENKTCITIVETRPSADFTFALLKPP--GSLIKHGSIP-----SDL----- 366  
QY 243 PVPVDVWMTARN-----ENQVVOPO---NGRVTLDGELLG-TTPLLAVNTCKPKGEVI 291  
Db 367 -TPRNSAHMGNRWSTISGFSVQPRVFSQNRHFDFTSTTGWSTPYYPYPIKIQGKVG 425  
QY 292 AKNGDVSRYMDMEITNDGTPTDPTEDTPG-PIGSPDFO-----GILFGVASOR 340  
Db 426 SNN-----KWFHVIDTDKALV-----PGPDGWPDTTIPDETAKTNGNFSYGESYRA 472  
QY 341 KNNEQNPAATRAEAIINTGGDHLCPQISSSEI 372  
Db 473 GSTTIKP-----NENSTHFKGTGICGLSTVEI 500

RESULT 10  
VCWHP9  
coat protein - feline calicivirus (strain F9)  
N:Alternate names: capsid protein

C:Species: feline calicivirus  
C>Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #text\_change 28-Jul-2000  
C:Accession: B43382; C45538; PQ0407; S23702  
R:Carter, M.J.; Milton, I.D.; Meanger, J.; Bennett, M.; Gaskell, R.M.; Turner, P.C.  
Virology 190, 443-448, 1992  
A:Title: The complete nucleotide sequence of a feline calicivirus.  
A:Reference number: A43382; MUID:92410623; PMID:1529544  
A:Accession: B43382  
A:Molecule type: genomic RNA  
A:Residues: 1-671 <CAR1>  
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
R:Carter, M.J.; Milton, I.D.; Turner, P.C.; Meanger, J.; Bennett, M.; Gaskell, R.M.  
Arch. Virol. 122, 223-235, 1992  
A:Title: Identification and sequence determination of the capsid protein gene of feline calicivirus.  
A:Reference number: A45538; MUID:92117861; PMID:1731695  
A:Accession: C45538  
A:Molecule type: Genomic RNA; protein  
A:Residues: 1-671 <CAR2>  
A:Cross-references: GB:M86379; NID:g323877; PIDN:AAA79327.1; PID:g323879  
A:Experimental source: strain F9  
A:Note: sequence extracted from NCBI backbone (NCBIN:77457, NCBIP:77462)  
R:Guiver, M.; Littler, E.; Caul, E.O.; Fox, A.J.  
J. Gen. Virol. 73, 2429-2433, 1992  
A:Title: The cloning, sequencing and expression of a major antigenic region from the feline calicivirus.  
A:Reference number: PQ0407; MUID:93019069; PMID:1402818  
A:Accession: PQ0407  
A>Status: nucleic acid sequence not shown  
A:Molecule type: mRNA  
A:Residues: 217-266, 'SI', 269-336, 'V', 338-395, 'S', 397-412, 'A', 414-521, 'EV', 524-671 <GUI>  
A:Cross-references: PIDN:AAB23553.1; PID:g257083  
C:Superfamily: feline calicivirus coat protein  
C:Keywords: capsid protein; coat protein; glycoprotein  
F:177,304,439,459,618/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 9.1%; Score 262.5; DB 1; Length 671;  
Best Local Similarity 25.3%; Pred. No. 4.3e-12;  
Matches 113; Conservative 59; Mismatches 167; Indels 107; Gaps 20;

QY 30 EPVAGASLATPVVGGQNIIDPWIRNPFVQAPAGEFTV---SPRNSPGEILLDLELGLDLP 86  
Db 145 EPSAQWSTAADMATKCSVDSEW-----EAPFSPTSVNWSSTETQKILFKQSLGPLN 198  
QY 87 PYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIIFAAIPPGF-PYENLSPSQITMCPHVI 145  
Db 199 PYLEHLAKLYVAVSGSIEVRFISGSGVFGKLAALVVPVPGVDPVQSTSMLOQ---PHVL 255  
QY 146 IDVROLEPFLMPDIWNNFFHYNQGNPKLRVLVAMLYTPLRANSGDDVFTVSC--RVL 203  
Db 256 FDARQVEPVIFCLPDLRSTLHLMSDTD-TTSLVIMVNDL-INPYANDANSGCCIVTVE 313  
QY 204 TKPSPDFEFTLVLPPTVESKTKQFALPILKISEMNSRFP---VP-----VDVMYTARN 254  
Db 314 TKPGDPFKFHLKPPG-----SMLTHGSIPTSLIPKTSLSLWIGNRWSDI 358  
QY 255 ENQVVOPO---ONGRVTLDGELLG-TTPL---LAVNICKPKGEVIAKNGDVSRYM----- 302  
Db 359 TDFVIRPFVQANRHFDENQETAGNSTPRFRDISVTITEQNGAKLG-IGVATDIYVPGIP 417  
QY 303 -----DMEITNDGTPTDPTDTE--DTPGPI-CSPDFOGLFGVASORKN 343  
Db 418 DGMWPTTIPGELIPAGDYAITNGDITATGYDTADIIKNNTNFRGMYICGSLQRAWG 477  
QY 344 EQNPATRAEAIINTGGD-----HLCPOQISSSEIYL-----TSPNLI----- 380  
Db 478 DKKISNTAFITATLDGNNKINFCNTIDQSKIIVFDQNHVGGKKAQTSDDTLAALGYTG 537  
QY 381 -----RCTNPQPLPQSGLRG 395  
Db 538 IGEQAIGSDRDRVVRISTLPETGARG 563

RESULT 11  
A48562

145	DB	EP5AQMWSTAADWASGKSYDSEW-----EAPSFHTSVNWSSTSETQGKILFKQSLGPLLN 198
87	QY	PYLAHLARMYNGHAGGMEVOIVLAGNATAGKIIFAAIPPGF-PYENLSPSOITMCPHV 145
199	DB	PYLEHLKSLKYVWAGSGSIIVRFSGVFGGKLAIAIVPPGVDPVQOSTSMLQY----PHVL 255
146	QY	IDVROLEBPELLMPDIWNFFHYNQNDPDKLRLVAMLYTFL-----RANNSGGDDVFTVSC 200
256	DB	FDARQVEPVIFIPDLRSTLYHVMSDTD--TTSLVIMVYNDLINPYANDSNSSGCIIVTE- 313
201	QY	RVLTKPSDPDEFTEFLVPP----- 218
314	DB	--TKPGDPDFKPHLLKPPGCVLTHGSIIPSDLIPKSSLSLWIGNRYWTDITDFVIRPFVFOA 370
219	QY	-----TVESKTKQKQFALPILKISSEMTNSRFPVPDVMY-----TARNENQV 258
371	DB	NRHFDPNQETAGWSTPRFPITITISEKNGSKLGIGVATDYIIPGIPDGWPDTTIADKLI 430
259	QY	VQOPQNGRVTLDELGLGT-----TPLLAVNICKPKGEVIAKNGDVRSYRMDMEITNT---- 309
431	DB	PAGDYSITTTGEGNDIKTAQAYDTAAVVKNTTFRGMVIC--GSLQRAWGDKKISNTAFIT 488
310	QY	-----DGTPIPDPT 317
489	DB	TAIRDGNEIKPS 500
<p>RESULT 13</p> <p>VCMWFF</p> <p>coat protein - feline calicivirus (strain CFI/68 FIV)</p> <p>N:Alternate names: capsid protein</p> <p>C:Species: feline calicivirus</p> <p>C:Date: 30-Jun-1992 #sequence revision 30-Jun-1992 #text_change 20-Aug-1999</p> <p>C:Accession: A40507; B40507; T09246</p> <p>R:Neill, J.D.; Reardon, I.M.; Heinriksson, R.L.</p> <p>J. Virol. 65, 5440-5447, 1991</p> <p>A:Title: Nucleotide sequence and expression of the capsid protein gene of feline calicivirus</p> <p>A:Reference number: A40507; MUID:91374597; PMID:1716692</p> <p>A:Accession: A40507</p> <p>A:Molecule type: genomic RNA</p> <p>A:Residues: 1-668 &lt;NEI&gt;</p> <p>A:Cross-references: GB:M32819; NID:G323874; PIDN:AAA42925.1; PID:G323875</p> <p>A:Accession: B40507</p> <p>A:Molecule type: protein</p> <p>A:Residues: 373-379; 403-419; 481-489; 560-566 &lt;NE2&gt;</p> <p>R:Neill, J.D.</p> <p>submitted to the EMBL Data Library, April 1998</p> <p>A:Description: Complete nucleotide sequence of feline calicivirus strain CFI/68</p> <p>A:Reference number: Z16626</p> <p>A:Accession: T09246</p> <p>A:Status: preliminary; translated from GB/EMBL/DBD</p> <p>A:Molecule type: genomic RNA</p> <p>A:Residues: 1-668 &lt;NE3&gt;</p> <p>A:Cross-references: EMBL:U13992; NID:G3056875; PIDN:AAAC13993.1; PID:G537256</p> <p>A:Experimental source: strain CFI/68 FIV</p> <p>C:Superfamily: feline calicivirus coat protein</p> <p>C:Keywords: capsid protein; coat protein; glycoprotein</p> <p>F:177,301,304,439,459,615/Binding site: carbohydrate (Asn) (covalent) #status F</p> <p>Query Match 8.0%; Score 254; DB 1; Length 668;</p> <p>Best Local Similarity 36.7%; Pred. No. 1.9e-11;</p> <p>Matches 58; Conservative 26; Mismatches 60; Indels 14; Gaps 5</p>		
QY	67	SPRNSGCEILLDLGLDNLNPLYLAHLARMYNGHAGGMEVOIVLAGNATAGKIIFAAIP 126
DB	179	STSETQKILFKQSLGPLNPNYIHLAKLVANGSGVDVRFSGVFGGKLAIAIVVP 238
QY	127	GF-PYENLSPSQITMCPHVIIIDVROLEBPELLMPDIWNFFHYNQNDPKLRLVAMLYTP 185
DB	239	GIDPVQOSTSMLQY---PHVLFARQVEPVIFISIPDLRSTLYHLMSTDID--TTSLVIMVYND 294
QY	186	L-----RANNSGGDDVFTVSCRVLTKPSDPDEFTEFLVPP 218



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:34:56 ; Search time 6.53557 Seconds  
(without alignments)  
4310.252 Million cell updates/sec

Title: US-09-926-799-11  
Perfect score: 2894  
Sequence: 1 MKWASNDAPSDGAGLVLP.....VNQFYS LAPVTGKGRRVQ 541

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1563	54.0	539	1 COAT_LORDV	P54635 lordsdale v
2	966	33.4	546	1 COAT_SOUV3	Q04542 southampton
3	311	10.7	2344	1 POLN_RHDV	P27410 rabbit hemo
4	271.5	9.4	703	1 COAT_SMSV4	P36285 san miguel
5	262.5	9.1	671	1 COAT_FCVF9	P27406 feline cali
6	262	9.1	702	1 COAT_SMSV1	P36284 san miguel
7	260	9.0	668	1 COAT_FCVF4	P27405 feline cali
8	254	8.8	668	1 COAT_FCVF6	P27404 feline cali
9	247.5	8.6	2208	1 POLN_MANCV	Q69014 manchester
10	126.5	4.4	2194	1 POLG_HE701	P32537 h genome po
11	125.5	4.3	2193	1 POLG_CX16T	Q9qf31 c genome po
12	124.5	4.3	2193	1 POLG_CX16G	Q65900 c genome po
13	121.5	4.2	2184	1 POLG_EC01F	Q91734 e genome po
14	120	4.1	2193	1 POLG_HE71M	Q66479 h genome po
15	118.5	4.1	2233	1 POLG_FMDV1	P03306 f genome po
16	118	4.1	2183	1 POLG_CXB4E	Q86887 c genome po
17	118	4.1	2193	1 POLG_HE71B	Q66478 h genome po
18	117.5	4.1	2205	1 POLG_POL2W	P23069 p genome po
19	117	4.0	2175	1 POLG_BOVEV	P12315 b genome po
20	115.5	4.0	2206	1 POLG_POL1M	P03299 p genome po
21	115.5	4.0	2208	1 POLH_POL1M	P03300 p genome po
22	113.5	3.9	2206	1 POLG_CXA21	P22055 c genome po
23	112.5	3.9	2303	1 POLG_TWEVG	P08545 t genome po
24	112	3.9	2185	1 POLG_SVDVH	P16504 a genome po
25	112	3.9	2301	1 POLG_TNEVD	P13899 t genome po
26	112	3.9	2332	1 POLG_FMDVA	P03308 f genome po
27	111	3.8	2183	1 POLG_CXB4J	P08292 c genome po
28	110.5	3.8	862	1 POLG_EC16H	Q66790 echovirus 1
29	110	3.8	2209	1 POLG_POL1S	P03301 p genome po
30	109	3.8	2185	1 POLG_SVDVU	P13900 s genome po
31	109	3.8	2290	1 POLG_EMCV	P03304 e genome po
32	108.5	3.7	2207	1 POLG_POL2L	P06210 p genome po
33	106.5	3.7	2194	1 POLG_EC30B	Q9wn78 e genome po

ALIGNMENTS

RESULT 1

ID	COAT_LORDV	STANDARD;	PRT;	539 AA.
AC	P54635;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	01-NOV-1997 (Rel. 35, Last annotation update)			
DE	Coat protein (Capsid protein).			
OS	Lordsdale virus (Human enteric calicivirus).			
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;			
OC	Norovirus.			
OX	NCBI_TaxID=82658;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96005060; PubMed=7561776;			
RA	Dingle K.E., Lambden P.R., Caul E.O., Clarke I.N.;			
RT	"Human enteric Caliciviridae: the complete genome sequence and expression of virus-like particles from a genetic group II small round structured virus."			
RL	GenBank:U00004.4 (1983-1984)			
CC	-1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.			

CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).
CC	EMBL; X86557; CAA60295.1; InterPro; IPR004005; Calici_coat.
DR	InterPro; IPR008975; Viral_Cap_coat.
DR	Pfam; PF00915; Calici_coat; 1.
KW	Coat protein; Glycoprotein.
FT	CARBOHYD 172 172 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 307 307 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 406 406 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT	CARBOHYD 531 531 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ	SEQUENCE 539 AA; 58775 MW; 5E5C63E7F2C5FD21 CRC64;

Query Match	54.0%; Score 1563; DB 1; Length 539;
Best Local Similarity	54.3%; Pred. No. 4.7e-111;
Matches 305; Conservative	76; Mismatches 133; Indels 48; Gaps 8;
QY	1 MKWASNDAPSDGAGLVLPETNNVMPLEPVAGASLATPVVGGQNIIDPWNRNPFVQAP 60
Db	1 MKWASNDAPSDGSANLVPEVNNVMALEPVVGAIAAPVAGQNVIDPWNRNPFVQAP 60
QY	61 AGEFTVSPRNSPEILLDLLEGLPDLNPLAHLARMYNGHAGMEVQIVLAGNAFTAGKTI 120
Db	61 GGEFTVSPRNPGEILWSAPGLPDLNPLVSLSRMYNGVAGGFQVILAGNAFTAGKVI 120
QY	121 FAATPGPGPYENLSQITMCHVIDVROLEPFLIPMPDDIWNFFHYNQNDPKRLVA 180
Db	121 FAAPVPNFFTEGLSPSQVTMPFHHIIVDRQLEPVLIPDPVRNFFHYNQANDSTLKLTA 180

```

QY 181 MLYTLRANNSGDDVFTVSCRVLTKPSPDFEFVLVPTVSKTKQFALPILKISEMTNS 240
DB 181 MLYTLRANNSGDDVFTVSCRVLTKPSPDFEFVLVPTVSKTKQFALPILKISEMTNS 240
QY 241 RFPVVDVVMYTARNENQVQVQNGRVTLDGELLGTTPLLAIVNICKFKGVIKNGDVRSY 300
DB 241 RFPVVDVVMYTARNENQVQVQNGRVTLDGELLGTTPLLAIVNICKFKGVIKNGDVRSY 300
QY 301 RMDMEITNDGTPIDPTDTPGIGSPDFQGLIFGVASORNNKNEQNAPTRAHEAINTGG 360
DB 298 DVTNMLASQWNSYDTEIPAPLPGTDFVGKIQGLLTQTR--ADGSTRAHKATVSTGS 355
QY 361 DHLCPQISSEIYLTSPNLRCTNPQLPQSGLRGTILIRSDNGCHDMVGTSPITPTWP 420
DB 356 VHTFKLGSVQFTDNTNDFQAGQNTKFTFVG-----VIQDGDHONE-----P 399
QY 421 QQWRRCRSGNCCSGHRYPP--VPVVMNRVTWVLSHKSGFSTSRKLPQLNL-----R 472
DB 400 QQWLSLNYSGRTCHNVHLAPAVAPTFFGQLLFFRSTMFQCS-----GYFNMLDCLLPQ 455
QY 473 W-----PLIRFNPDTGRVLFEARLHKQGITVAHTGDNPIVMVNGYPR 517
DB 456 WVLHFYQEAAPQSDVALLRFVNPDTGRVLFECKLHKSGYITVAHTGPDYDLVLPNGYPR 515
QY 518 FEAWNQFYSLAPVGTGKRRR 539
DB 516 FDSVWNQFYTLPAMGNGTGRRR 537

RESULT 2
COAT_SOUV3 STANDARD; PRT; 546 AA.
ID COAT_SOUV3
AC Q04542;
DT 01-OCT-1994 (Rel. 30, Created)
DT 01-OCT-1994 (Rel. 30, Last sequence update)
DE Coats protein (Capsid protein).
OS Southampton virus (serotype 3).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=37129;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-93142023; PubMed-8380940;
RA Lambden P.R., Caul E.O., Ashley C.R., Clarke I.N.;
RT "Sequence and genome organization of a human small round-structured (Norwalk-like) virus.";
RL Science 259:516-519(1993).
CC -1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; L07418; AAA92984.1; -.
CC PIR; B37491; B37491.
CC DR InterPro; IPR004005; Calici_coat.
CC DR InterPro; IPR008975; Viral_cap_coat.
CC DR Pfam; PF00915; Calici_coat; 1.
CC Coats protein; Glycoprotein.
CC CARBOHYD 73 73 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 303 303 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 340 340 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC CARBOHYD 441 441 N-LINKED (GLCNAC. . .) (POTENTIAL).
CC SEQUENCE 546 AA; 58774 MW; C82B2A85AD4B05EA CRC64;
Query Match 33.4%; Score 966; DB 1; Length 546;
Best Local Similarity 38.5%; Pred. No. 9.8e-66;

```

```

Matches 225; Conservative 87; Mismatches 189; Indels 84; Gaps 17;
QY 1 MKMASNDAAPSDCAAG--LVPEINN-EVMPLEPVAGASLATPVVGQNIIDPIWRNFF 56
DB 1 MKMASNDAAPSDCAAG--LVPEINN-EVMPLEPVAGASLATPVVGQNIIDPIWRNFF 56
QY 57 VQAPAGETVSPRNSPGEIILDLGLPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAPTA 116
DB 61 VQSPQGETIISNTPTPGDILFDLQGLPHLNPFLSHLSQMYNGWGNMVRVILLAGNAPSA 120
QY 117 GKILPAALPPGPPVENISPSQITMCPHVIDVROLEPFLPMPDIWNFFHYVNGNDPKL 176
DB 121 GKIIIVCCVPPGFTSSSLTIAQTALFPHVADVTRLEPIEMPLEDVRNVLYHTND-NQPTM 179
QY 177 RLVAMLYTLPLR--ANNSGDDVFTVSCRVLTKPSPDFEFVLVPTVSKTKQFALPILKI 234
DB 180 RLVCMLYTLPLRTGGSGNSDSFVAGRVLTAPSDFSFLVLPPTIEQKTRAFVNPNIPL 239
QY 235 SEMTNSRFPVVDVVMYTARNENQVQVQNGRVTLDGELLGTTPLLAIVNICKFKGVIK 294
DB 240 QTLNSRFPVVDVVMYTARNENQVQVQNGRVTLDGELLGTTPLLAIVNICKFKGVIK 296
QY 295 GDVRSYRMDMEITNDGTPIDPTDTPGIGSPDFQGLIFGVASORNNKNEQNAPTRAHEA 354
DB 297 QGART-----LNLTEVDGKPP--MAFDSPAPVGPDPFGKCDMWMRISKTKPNTGSGDPMRSV 351
QY 355 IINTGGDHLCPQISS--SEIYLTSPNLRCTNPQLPQSGLRGTILIRSDNGCHDMVG 411
DB 352 SVQTNVQGVFHLGSIQFDEVF-----NHPTGDYIG 382
QY 412 T-----SPTTP--TWPOQWRRCSRGNCCSGHRYPPVPMVNRVTWVLSHKSGFSTSTR 464
DB 383 TIEMISQPTSPGTDINLWEIPDYGSSLSQAANLAP-PVFPFPGFGEALVYFVSAPPGPN 441
QY 465 K-----LPO-----LNLRW-----LIRFNPDTGRVLFEARLHKQGITVA-- 501
DB 442 RSAPNDVPCLLPQEIYTHFVSEQATMGDAALLHYVDPDTRNLGFEFLPGGYLTCVNP 501
QY 502 HTGDNPIVMVNGYPRFEAWNQFYSLAPVGTGK-----GRRRV 540
DB 502 GVGAGPQQLPLNGVFLFVSWSRFQLKAPVGTASTARGSLGVRI 546

RESULT 3
POLN_RHDV STANDARD; PRT; 2344 AA.
ID POLN_RHDV
AC P27410;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE Non-structural polyprotein [Contains: RNA-directed RNA polymerase (EC 2.7.7.48); Thiol protease P3C (EC 3.4.22.-); Helicase (2C like protein); Coat protein].
OS Rabbit hemorrhagic disease virus (RHDV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Lagovirus.
OX NCBI_TaxID=11976;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE-91361557; PubMed=1840711;
RA Meyers G., Wirblich C., Thiel H.-J.;
RT "Rabbit hemorrhagic disease virus -- molecular cloning and nucleotide sequencing of a calicivirus genome.";
RL Virology 184:664-676(1991).
CC -1- FUNCTION: THE P3C CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE
CC FOR CLEAVAGE AT CERTAIN Q/G SITES IN THE POLYPROTEIN.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
CC {RNA} (N).
CC -1- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its

```



```

Db 426 SNN-----KHFVHIDTKALV-----PGIPDGHPTDITPDKATNGNSYGESYRA 472
QY 341 NKNEQNQPATRAHEAIINTGDLCPQISSEI 372
Db 473 GSTTIKP-----NENSTHFKGTICGLTSTVEI 500

RESULT 5
COAT_FCVF9
ID COAT_FCVF9 STANDARD; PRT; 671 AA.
AC P27406;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DE Coate protein (Capsid protein).
GN Cpp76.
OS Feline calicivirus (strain F9) (FCV).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=11981;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92410623; PubMed=1529544;
RA Carter M.J., Milton I.D., Meanger J., Bennett M., Gaskell R.M.,
RA Turner P.C.;
RT "The complete nucleotide sequence of a feline calicivirus.";
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92339673; PubMed=1633955;
RA Meanger J., Carter M.J., Gaskell R.M., Turner P.C.;
RT "Cloning and sequence determination of the feline calicivirus strain
RT F9.";
RL Biochem. Soc. Trans. 20:26S-26S(1992).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M86379; AAA79327.1; -.
DR EMBL; Z11536; CAA77636.1; -.
DR PIR; B43382; VCWVF9.
DR InterPro; IPR004005; Calici coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 177 177 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 304 304 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 439 439 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 459 459 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 618 618 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 671 671 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 671 AA; 73441 MW; 33BE86D8370D5E5 CRC64;

Query Match 9.1%; Score 262.5; DB 1; Length 671;
Best Local Similarity 25.3%; Pred. No. 3.2e-12;
Matches 113; Conservative 59; Mismatches 167; Indels 107; Gaps 20;

QY 30 EPVAGASLAPTVGQQNIIDPWIRNNFVQAPAGEFTV---SPRNSPGEILLDLLELGPDLN 86
Db 145 EPSAQMSAADMATGKSGVDSEW-----EAFSFTSVNWNSTSETQKILFKQSLGPLLN 198
QY 87 PYLAHLARMYNGHAGGVEOIVLAGNAFTAGKIIFRAIPPGF-PYENLSPSQITMCPHVI 145
Db 199 PYLEHLAKLVANSGSTVEAFSGSGVFGKLAIVVPGVDPVQSTMLQY---PHVL 255
QY 146 IDVRQLEPFLPMPDIWNFFHYNQGNPDKRLRLVAMLYTFLRANNSGDDVFTVSC--RVL 203

```

---

```

Db 256 FDARQVEPIFCLPDLARSTLYHLMSDITD--TTSLVIMVYNDL-INPYANDANSSGCIPTVE 313
QY 204 TKPSPDFEPTLVPTVESKTKQFALPILKISEMTNSRFP---VP-----VDMVYTARN 254
Db 314 TKPGPDFKPHLLKPPG-----SMUTHGSIPSDLLPKTKTSSLWIGHNRYWSDI 358
QY 255 ENQVVQP-----QNGRVTLIDGELLG--TTPL-----LAVNICKFKFGEVIKAGNDVRSYRM----- 302
Db 359 TDFVIRPFVFOANRHFDFNQETAGWSIPRFRPISVITTEQNGAKLG-IGVATDIYVPGIP 417
QY 303 -----DMEITNTDGTPIIDPTE--DTFGPI--GSPDFQGLIFGVASORKN 343
Db 418 DGWPDTTIPGELLIPAGDYAITNGTNDITATGYDTADIIKNTNFRGMYTCGSLQRAWG 477
QY 344 EQNPATRAHEALINTGDD-----HLCPQISSSEIVL-----TSPNIL----- 380
Db 478 DKISNTAFITATLDGNNKINPCNTIDOSKIVVFDQNHVHGKKAQTSDDTLALLGYTG 537
QY 381 -----RCTNPQPLQSGLRG 395
Db 538 IGEQAIGSDRDRVVRVISTLPETGARG 563

RESULT 6
COAT_SMSV1
ID COAT_SMSV1 STANDARD; PRT; 702 AA.
AC P36284;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-JUN-1994 (Rel. 29, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Coat protein (Capsid protein).
OS San Miguel sea lion virus (serotype 1) (SMSV 1).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Vesivirus.
OX NCBI_TaxID=36406;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92410750; PubMed=1529644;
RA Neill J.D.;
RT "Nucleotide sequence of the capsid protein gene of two serotypes of
RT San Miguel sea lion virus: identification of conserved and non-
RT conserved amino acid sequences among calicivirus capsid proteins.";
RL Virus Res. 24:211-222(1992).
CC -!- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; M87481; AAA16217.1; -.
DR PIR; A48562; A48562.
DR InterPro; IPR004005; Calici coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
KW Coat protein; Glycoprotein.
FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 481 481 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 493 493 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 545 545 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 702 AA; 77850 MW; E6E5A58523DEE3D7 CRC64;

Query Match 9.1%; Score 262; DB 1; Length 702;
Best Local Similarity 27.2%; Pred. No. 3.7e-12;
Matches 96; Conservative 52; Mismatches 125; Indels 80; Gaps 19;

QY 8 AAPSSDGAAGLYPEINNEVMPLEPVAGASLATPVVVGQQNIIDPWIRNNFVQAPAG----- 62
Db 151 AESDGGPGGADIVTEEGTVVQQQPVPAQSAITLAA-----ASTGKTVDVC 195

```



QY 63 EETV-----SPRSPGEILLDLGPDLPNLYLAHARMYNGHAGMEVOIVLAGN 112  
 Db 196 EMTTFSYHTAVNMSTTEAQKILFSRALSPNLYLRHISLSYSTWGGIDVRPTVSGS 255  
 QY 113 AFTAGKIIFAAIPPGF-PYENLSPSOITWCPHVIIIDVROLEPFLPMPDIWNFFHYNOG 171  
 Db 256 GVFGKLAALIVPPGLEVE--SPTML-QYPHVLFDARQTEPVIFPIPIRKLTHSMDD 312  
 QY 172 NDKRLVLAULTYPL-RANNSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALP 230  
 Db 313 TD-TTRLVIMVYNELINPEQSEPKSSCSITVETRPSSDFTSLLKPP--GSLKHGSI 369  
 QY 231 ILKISEWTSRPPVPVDMYTARN-----ENQVQPOQ-----NGRVLGELLG-TTPL- 278  
 Db 370 ----SOL-----IPNSRHMGNRWMSITDFVQVQFQSNRHPDFDSTTTGWSPTY 419  
 QY 279 --LAVNICKFKGEVIAKNGDVSRYRMDMEITNDGTPIDPTEDTPG-PIGSPD 328  
 Db 420 IEIVTL-----EKLDGGQY-----FKVTDEKSLV-----PGLPDGWF 455  
 RESULT 7  
 COAT\_FCVF4  
 ID COAT\_FCVF4 STANDARD; PRT; 668 AA.  
 AC P27405;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 01-JUN-1994 (Rel. 29, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS Feline calicivirus (strain Japanese F4) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OC NCBI\_TaxID=11980;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91306470; PubMed=1853578;  
 RA Tonya Y., Taniguchi Y., Takahashi E., Utagawa E., Takeda N.,  
 Miyamura K., Yamazaki S., Mikami T.;  
 RT "Sequence analysis of the 3'-end of feline calicivirus genome.";  
 RL Virology 183:810-814(1991).  
 CC 1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; D90357; BAAL4371.1; -.  
 DR PIR; B40481; VCMWFC.  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 399 399 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 668 AA; 73589 MW; 85BDBC85904E503 CRC64;  
 Query Match 9.0%; Score 260; DB 1; Length 668;  
 Best Local Similarity 25.0%; Pred. No. 4.9e-12;  
 Matches 93; Conservative 46; Mismatches 133; Indels 100; Gaps 12;  
 QY 30 EPVAGSLATPVGOQNIIDPWIRNFFVQAPAGEFTV---SPRSPGEILLDLGPDLPN 86  
 Db 145 EFSAQNSTAADNASKGVDSW-----EAFSFTSVNWSSTQGGKILFKQSLGPLLN 198  
 QY 87 PYLAHARMYNGHAGMEVOIVLAGNAFTAGKIIFAAIPPGF-PYENLSPSOITWCPHY 145

Db 199 PYLEHLSKLYVAMSGSIEVRFSGVGGKLAIVVPPGVDPQOSTMLOY----PHVL 255  
 QY 146 IDVROLEPFLPMPDIWNFFHYNOGNDPKLRLVLAULTYPL-----RANNSGDDVFTVSC 200  
 Db 256 FPARQVEPVIFPIPIRLSTLYHVMSTD-TTSLVIMVYNELINPYANDSNSSGCIITVE- 313  
 QY 201 RVLTSPSDPFEFTFLVPP----- 218  
 Db 314 ---TKPGDPKFHLKPPGSLVTHGSIIPSDLIPIKSSSLMIGNRYWYTDITDFVIRPFFQA 370  
 QY 219 -----TVESKTKQFALPILKISEMNTSRFPVDMY-----TARNENOV 258  
 Db 371 NRHPDFOETAGSTPRFRPITITISEKNGSKLGIGVATDYIIPGIPDGWPTTTADKLI 430  
 QY 259 VQOPNGRVTLDGELLGT-----TPLLAVNICKFKGEVIAKNGDVSRYRMDMEITNT 309  
 Db 431 PAGDYSIITGEGNDIKTAQAYDTAAVKNVNTNFRGMWYC--GSLQRAWGCDKISNTAFIT 488  
 QY 310 ----DGTPIDPT 317  
 Db 489 TAIRDGNEIKPS 500  
 RESULT 8  
 COAT\_FCVC6  
 ID COAT\_FCVC6 STANDARD; PRT; 668 AA.  
 AC P27404;  
 DT 01-AUG-1992 (Rel. 23, Created)  
 DT 01-AUG-1992 (Rel. 23, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Coat protein (Capsid protein).  
 OS Feline calicivirus (strain CFI/68 FIV) (FCV).  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Vesivirus.  
 OC NCBI\_TaxID=11979;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=91374597; PubMed=1716692;  
 RA Neill J.D., Reardon I.M., Heinrichson R.L.;  
 RT "Nucleotide sequence and expression of the capsid protein gene of  
 RT feline calicivirus";  
 RL J. Virol. 65:5440-5447(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Neill J.D.;  
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.  
 CC 1- PTM: The N-terminus is blocked.  
 CC 1- SIMILARITY: TO CAPSID PROTEIN OF OTHER CALICIVIRUSES.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; M32819; AAA42925.1; -.  
 DR EMBL; U13992; AAC13993.1; -.  
 DR PIR; A40507; VCMWFF.  
 DR InterPro; IPR008975; Viral\_cap\_coat.  
 DR InterPro; IPR00915; Calici\_coat; 1.  
 KW Coat protein; Glycoprotein.  
 FT CARBOHYD 177 177 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 301 301 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 304 304 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 439 439 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 459 459 N-LINKED (GLCNAC. .) (POTENTIAL).  
 FT CARBOHYD 615 615 N-LINKED (GLCNAC. .) (POTENTIAL).  
 SQ SEQUENCE 668 AA; 73550 MW; 9E52312108D441 CRC64;

Query Match 8.8%; Score 254; DB 1; Length 668;  
Best Local Similarity 36.7%; Pred. No. 1.4e-11; Mismatches 60; Indels 14; Gaps 5;  
Matches 58; Conservative 26;

QY 67 SPNSGGEILLDLGLPDLNPLYLAHLARMYNGHAGMEYQIVLAGNAFTAGKIIFAAIAPP 126  
DB 179 STSETQKGLFKQSLGFLNPLYLTHLAKLYVAWSGVDRFSISGSGVFGKLAIVVPP 238  
QY 127 GF-PYENLSPSQITMCPHVIDVRQLEPFLPMPDIWNNFFHYNQNDPKLVLVAMLYTP 185  
DB 239 GIDPQVSTMLQY---PHVLFDAQVERPVIIFSIPDLRSTLYHLSMTD-TTSLVMVND 294  
QY 186 L-----RANSGDDVFTVSCRVLTKPSPDFEFTFLVPP 218  
DB 295 LINPYANDSSGCVITVE-----TKGPDFKFHLKPP 328

RESULT 9  
POLN\_MANCV STANDARD; PRT; 2208 AA.  
ID POLN\_MANCV STANDARD; PRT; 2208 AA.  
AC Q69014;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Genome polyprotein [Contains: RNA-directed RNA polymerase  
(EC 2.7.7.48); Thiol protease 3C (EC 3.4.22.-); Helicase (2C like  
protein); Coat protein].  
OS Manchester virus (Human enteric calicivirus).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Sapovirus.  
OX NCBI\_TaxID=82659;  
RN [1]  
SEQUENCE FROM N.A.  
RX MEDLINE=95390791; PubMed=7661689;  
RA Liu B.L., Clarke I.N., Caul E.O., Lambden P.R.;  
RT "Human enteric caliciviruses have a unique genome structure and are  
distinct from the Norwalk-like viruses.";  
RL Arch. Virol. 140:1345-1356(1995).  
CC -!- FUNCTION: P2C IS IMPORTANT IN RNA REPLICATION (BY SIMILARITY).  
CC -!- FUNCTION: THE CYSTEINE PROTEASE IS THE PROTEASE RESPONSIBLE FOR  
THE POST-TRANSLATIONAL PROCESSING OF THE POLYPEPTIDE (BY  
SIMILARITY).  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
{RNA}(N).  
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC -!- SIMILARITY: TO PICORNAVIRUS POLYPEPTIDES.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C24.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----  
EMBL; X86560; CNA60262.1; --  
DR MEROPS; C24.0PW; --  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR004004; Calici\_pol\_hel.  
DR InterPro; IPR000317; Peptidase\_C24.  
DR InterPro; IPR000605; RNA\_helicase.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR InterPro; IPR007094; RNA\_pol\_PSVir.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
DR Pfam; PF03510; Peptidase\_C24; 1.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
DR PRINTS; PR00916; 2CENODPTASE.  
DR PRINTS; PR00918; CALICIVIRUSNS.  
KW Polyprotein; Transferase; RNA-directed RNA polymerase; Hydrolase;

KW Thiol protease; Helicase; ATP-binding; Coat protein.  
FT CHAIN ? ? HELICASE (P2C-LIKE).  
FT CHAIN ? ? PROTEASE (P3C).  
FT CHAIN ? ? RNA-DIRECTED RNA POLYMERASE.  
FT CHAIN ? ? COAT PROTEIN.  
FT NP\_BIND 408 415 ATP (POTENTIAL).  
FT ACT\_SITE 1097 1097 PROTEASE (BY SIMILARITY).  
FT ACT\_SITE 1112 1112 PROTEASE (BY SIMILARITY).  
SQ SEQUENCE 2208 AA; 242736 MW; 3E299D5BA20E45DC CRC64;  
Query Match 8.6%; Score 247.5; DB 1; Length 2208;  
Best Local Similarity 32.4%; Pred. No. 2.2e-10;  
Matches 73; Conservative 31; Mismatches 104; Indels 17; Gaps 6;  
QY 2 KMASNDAAAPSDGAAGLVPEINNEVM--PLEPVAGA---SLATPVVGQNIIDPWIRNRF 56  
DB 1658 KQSNPMVVDPPGTG--PTTSHVVVANPEQPGNAAQOELEAVATGAIQSNVPEAIRNCF 1715  
QY 57 VQAPAGEFTVSPRNSPGEILLDLGLPDLNPLYLAHLARMYNGHAGMEYQIVLAGNAFTA 116  
DB 1716 --AVPRTFAWDRMPTGTGFLGSLIHPNINPYTSHLSGMWAGWGSFEVRLSISGSGVFA 1773  
QY 117 GKILPAALPPGEPYENLSPSOIT---MCPHVIDVRQLEPFLPMPDIWNNFFHYNQND 173  
DB 1774 GRITASVTPPG-----VDPSSIRDPGVLPHAFVDAITEPVPVFMIPDVADVHYHRMOGAE 1828  
QY 174 PKLRLVAMLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPP 218  
DB 1829 PTCSLGFVWYQPLNPFSTTAVSCWVSVEYTKPGGDFDFCLLRPP 1873

RESULT 10  
POLG\_HE701 STANDARD; PRT; 2194 AA.  
ID POLG\_HE701 STANDARD; PRT; 2194 AA.  
AC P32537;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Genome polyprotein [Contains: Coat protein VP4 (PIA); Coat protein VP2  
(PIB); Coat protein VP3 (PIC); Coat protein VP1 (PID); Core protein  
P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-  
linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)  
(P3C); RNA-directed RNA polymerase P3D (strain J670/71) (EV 70).  
DE Human enterovirus 70 (strain J670/71) (EV 70).  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;  
OC Enterovirus.  
OX NCBI\_TaxID=31915;  
RN [1]  
SEQUENCE FROM N.A.  
RX MEDLINE=91037960; PubMed=2172447;  
RA Ryan M.D., Jenkins O., Hughes P.J., Brown A., Knowles N.J., Booth D.,  
RA Minor P.D., Almond J.W.;  
RT "The complete nucleotide sequence of enterovirus type 70:  
relationships with other members of the picornaviridae.";  
RT J. Gen. Virol. 71:2291-2299(1990).  
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN  
Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.  
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the  
poliovirus polypeptide. In other picornavirus reactions Glu may be  
substituted for Gln, and Ser or Thr for Gly.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +  
{RNA}(N).  
CC each of which is composed of one copy each of proteins VP1, VP2,  
VP3, and VP4.  
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>;  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

```
DR EMBL; D00820; BAAL1891.1; -.
DR PIR; A36253; GNNYE7.
DR HSSP; P03300; IPOV.
DR MEROPS; C03.UFA; -.
DR MEROPS; C03.UFB; -.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR004004; Calici_pol_hel.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000199; Pept_3C_picorn.
DR InterPro; IPR000081; Peptidase_C3.
DR InterPro; IPR003138; Pico_P2B.
DR InterPro; IPR002527; Pico_P2B.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR00605; RNA helicase.
DR InterPro; IPR007095; RNA_pol_DS_PS.
DR InterPro; IPR001205; RNA_pol_P3D.
DR InterPro; IPR007094; RNA_pol_PSVir.
DR Pfam; PF00548; Cys-protease_3C; 1.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00947; Pico_P2A; 1.
DR Pfam; PF01552; Pico_P2B; 1.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA_dep_RNA_pol; 1.
DR Pfam; PF00910; RNA_helicase; 1.
DR PRINTS; PR00918; CALICIVIRUS.
DR ProDom; PD001125; Cys_protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
DR Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 319 COAT PROTEIN VP2.
FT CHAIN 320 561 COAT PROTEIN VP3.
FT CHAIN 562 871 COAT PROTEIN VP1.
FT CHAIN 872 1014 CORE PROTEIN P2A.
FT CHAIN 1015 1113 CORE PROTEIN P2B.
FT CHAIN 1114 1443 CORE PROTEIN P2C.
FT CHAIN 1444 1532 CORE PROTEIN P3A.
FT CHAIN 1533 1554 GENOME-LINKED PROTEIN VPG.
FT CHAIN 1555 1737 PICORNAIN 3C.
FT CHAIN 1738 2194 RNA-DIRECTED RNA POLYMERASE P3D.
FT LIPID 2 2 N-myristoyl glycine (in host) (By
similarity).
FT ACT SITE 1701 1701 PROTEASE (POTENTIAL).
FT ACT SITE 1715 1715 PROTEASE (POTENTIAL).
SQ SEQUENCE 2194 AA; 244590 MW; 15DBAE96EE06673C CRC64;

Query Match 4.4%; Score 126.5; DB 1; Length 2194;
Best Local Similarity 23.0%; Pred. No. 0.34;
Matches 65; Conservative 38; Mismatches 94; Indels 85; Gaps 13;

QY 21 EINNEWMLEPVAGASLAPVVGQQNIIDPWIRNNFVQAPAGEFTVSPRNSPGEILLDLLE 80
DB 373 EINN-----VNDASGVRLRVQISAQSDMDQLFN-----IPLDIQ 408

QY 81 L-GPOLNPYLAHLMRYNGHAGMEVQIIVLAGNAFTAGKIIIFAAIIPPGPPYENLSPS---- 136
DB 409 LSGPLRNTLLGNISRYTHWSGSLEWTFMFCGSPMTGKLIICYTPPG-----GSSPTDRM 464

QY 137 QITMCHVLIIDVRQLEPFLPMPDIWNFFHYNQGNPKLRLVAMLYPLRANNSGDDVVF 196
DB 465 QAMLATHVWVDFGLQSSITIIIP--WISGSHYR-----NFNTDAKAINA--NVG 509

QY 197 TVSCRVLTSPDFEFTFLVPT-----VSKTKQFALPILKISEMNTSPFPV 246
DB 510 YVTC-----FMQTNLVAVGADQCYIVGMVAACKDFNLRLMRDSP----- 550
```

```
QY 247 DVMYTARNENQVVPQNGRV-----TLDGEL---LGTTPLL 279
DB 551 DIGQSAILPEQAATTQIGIEIVKTVANTVESEIKAEGLVIPSL 592

RESULT 11
POLG_CX16T STANDARD; PRT; 2193 AA.
AC Q9QF31;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VPG (P3B); Picornain 3C (SC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Coxsackievirus A16 (strain Tainan/5079/98).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.
OC NCBI_TaxID=231417;
RX MEDLINE=21427116; PubMed=11536241;
RA Yan J.-J., Su I.-J., Chen P.-F., Liu C.-C., Yu C.-K., Wang J.-R.;
RT "Complete genome analysis of enterovirus 71 isolated from an outbreak in Taiwan and rapid identification of enterovirus 71 and coxsackievirus A16 by RT-PCR.";
RL J. Med. Virol. 65:331-339(2001).
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular membranes and is associated with viral RNA synthesis.
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-[Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
```

```
DR Pfam; PF00910; RNA helicase; 1.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR ProDom; PD001125; Cys protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
KW Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 323 COAT PROTEIN VP2.
FT CHAIN 324 565 COAT PROTEIN VP3.
FT CHAIN 566 862 COAT PROTEIN VP1.
FT CHAIN 863 1011 CORE PROTEIN P2A.
FT CHAIN 1013 1112 CORE PROTEIN P2B.
FT CHAIN 1112 1440 CORE PROTEIN P2C.
FT CHAIN 1441 1526 CORE PROTEIN P3A.
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VP6.
FT CHAIN 1549 1731 PICORNAIN 3C.
FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.
FT LIPID 2 2 N-myristoyl glycine (in host) (By
FT ACT_SITE 1695 1695 PROTEASE (POTENTIAL).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
FT SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;
Query Match 4.3%; Score 125.5; DB 1; Length 2193;
Best Local Similarity 22.1%; Pred. No. 0.4; Indels 93; Gaps 17;
Matches 88; Conservative 49; Mismatches 169;
SQ SEQUENCE 2193 AA; 243182 MW; 927839DB58F61E7F CRC64;
QY 1 MKMASNDAAAPSSDG-AAGLVP-----EINN--EVMPLPVPAGA-SLATPVVQ 44
Db 329 LKPGTNQFLTDDGVSAPIPLGPHPTPIHIGVEHVNLLICRVEITILEVNNLKT----- 383
QY 45 QNIIDPWIIRNF---VQAPAGFTVSPRNSPGEILLDLDELGPDLNLPYLALARMYNGHAG 101
Db 384 -NETTFMQRLCFVSVQSKTGELCAAFRADPGRD-----GPMQSTILGQLCRYTQWSG 436
QY 102 GNEVQIVLAGNAFTAGKIFAIAPGFPYENISPSQIT--MCPHVIIDVRQLEPFLPMP 159
Db 437 SLEVTFMFAGSMATGKMLIAIYTPPG---GNVPADRITAMLGHVTHVDFGLQSSVTLVVP 493
QY 160 DIWNFFHYNQGNQDPLRLVAMLYTPLRANNSGDDVFTVSCRVLTPSPDFFETFLVP-- 217
Db 494 -WISNTHYR-----AHARAGYFDVYTTGIITW-----YQTNVYVPIG 530
QY 218 -PT-----VESKTKQFALPILKISEMTNSRPV-----PVDVMYATARNENQV-----V 259
Db 531 APTTAYIVALAAQDNFTMKLCKDTEDEQTANIQGDPIADMIQDTVNNQVNRSLTALQV 590
QY 260 QPQNGRVTLGELLGTTPLLAVNICFKGEVIAKNGDVRSYRMDMEITWTDGTGTFID- 315
Db 591 LPTAADTEASSHRLGTVGVVPLQAQATGASSNASDKNLIETRCVLNHHSTQETAIGNFFS 650
QY 316 -----PTEDTPGPIGSPDFQGLFGVASQRKNKE 344
Db 651 RAGLVSIITMPTTGITQNTDGYVNWDDILNGYAQLRKCE 689
```

## RESULT 12

POLG\_CX16G

ID POLG\_CX16G STANDARD; PRT; 2193 AA.

AC Q65900;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2

DE (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein

DE P2A; Core protein P2B; Core protein P3A; Core protein P3A; Genome-

DE linked protein VP6 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C)

DE (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].

OS Coxsackievirus A16 (strain G-10).

OC Viruses; serRNA positive-strand viruses, no DNA stage; Picornaviridae;

```
OC Enterovirus.
OX NCBI_TaxID=69159;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=94303216; PubMed=8030260;
RX Poyry T., Hyypia T., Horsnell C., Kinnunen L., Hovi T., Stanway G.;
RT "Molecular analysis of coxsackievirus A16 reveals a new genetic group
of enteroviruses.";
RL Virology 202;982-987(1994).
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular
membranes and is associated with viral RNA synthesis.
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN
Q/G SITES IN THE POLYPEPTIDE. IT MAY BE A CYSTEINE PROTEASE.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the
poliovirus polyprotein. In other picornavirus reactions Glu may be
substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate +
[RNA] (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units,
each of which is composed of one copy each of proteins VP1, VP2,
VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL; U05876; AAA50478.1; -.
DR HSP; P03300; IPOV.
DR MEROPS; C03.022; -.
DR MEROPS; C03.UFA; -.
DR InterPro; IPR003593; AAA ATPase.
DR InterPro; IPR004004; Calici pol hel.
DR InterPro; IPR009003; Cys Ser trypsin.
DR InterPro; IPR000199; Pept_3C_picorn.
DR InterPro; IPR000081; Peptidase_C3.
DR InterPro; IPR003138; Pico_P1A.
DR InterPro; IPR002527; Pico_P2B.
DR InterPro; IPR001676; Rhv.
DR InterPro; IPR006005; RNA helicase.
DR InterPro; IPR007095; RNA pol DS PS.
DR InterPro; IPR001205; RNA pol_P3D.
DR InterPro; IPR007094; RNA pol_Psvir.
DR InterPro; IPR008975; Viral_cap_coat.
DR Pfam; PF00548; Cys-protease-3C; 1.
DR Pfam; PF02226; Pico_P1A; 1.
DR Pfam; PF00947; Pico_P2A; 1.
DR Pfam; PF01552; Pico_P2B; 1.
DR Pfam; PF00073; rhv; 3.
DR Pfam; PF00680; RNA dep RNA_pol; 1.
DR Pfam; PF00910; RNA helicase; 1.
DR PRINTS; PR00918; CALICIVIRUSNS.
DR ProDom; PD001125; Cys protease_3C; 1.
DR ProDom; PD001306; Pico_P2A; 1.
DR ProDom; PD001274; Pico_P2B; 1.
DR SMART; SM00382; AAA; 1.
DR Polyprotein; Coat protein; Core protein; Transferase;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;
KW Lipoprotein.
FT CHAIN 2 69 COAT PROTEIN VP4.
FT CHAIN 70 323 COAT PROTEIN VP2.
FT CHAIN 324 565 COAT PROTEIN VP3.
FT CHAIN 566 862 COAT PROTEIN VP1.
FT CHAIN 863 1011 CORE PROTEIN P2A.
FT CHAIN 1013 1112 CORE PROTEIN P2B.
FT CHAIN 1112 1440 CORE PROTEIN P2C.
FT CHAIN 1441 1526 CORE PROTEIN P3A.
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VP6.
```

```
FT CHAIN 1549 1731 PICORNAIN 3C.
FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE.
FT LIPID 2 N-methylglycine (in host) (By
FT similarity).
FT ACT SITE 1695 1695 PROTEASE (POTENTIAL).
FT ACT_SITE 1709 1709 PROTEASE (POTENTIAL).
SQ SEQUENCE 2193 AA; 243209 MW; 04B3BCE572A76E38 CRC64;

Query Match
Best Local Similarity 21.9%; DB 1; Length 2193;
Matches 86; Conservative 47; Mismatches 179; Indels 81; Gaps 15;

QY 1 MKVANSNDAPSSDG--AAGLVP-----EINNEWMPLEPVAGASLATPVG-QQNIIDP 50
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
329 LKPGTNQFLTTDDGVSAPILPGFHPPTPAIHGPEVRNLLIEICRVETILEVNNLSNETTP 388
QY 51 WTRNNF---VQAPAGEFTVSPNSGCEILLDLGLDLPNLYLAHLARMYNGHAGGMEVOI 107
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
389 MORLCFPVSVQSKTGELCAVFRADPG-----RNGPWQSTILGQLCRYTQWSGSLVTF 442
QY 108 VLAGNAFTAGKIFAAIPP--GPPYENLSPSQTTCMPHVIIDVRQLEPFLPMPDINWNNF 165
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
443 MPAGSFMATGKMLIATVPGGVGPADRLTAMLTG---HWIDFGLOSSVTLVIP--WISN 497
QY 166 FHYNQNDPKRLKLVAMLYTPLRANNSGDVFTVSCRVLTKPSDFEFTFLVP---PT--- 219
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
498 THYR-----AHAKQGYDYTTTGTITW-----YQTNVYWPIGAPTAY 536
QY 220 ---VSKTKQFALPKLISEMINSRFPVP-----VDVMTARNENQV-----VQPNQR 265
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
537 IVALAAQDNFTMKLCKDTEIQSANIQGDGIADIMDAQVTSRVRGALTSIQVEPTAAN 596
QY 266 VTLDGELLGTTLLAVNICKEVIAKNQGVRSYRMDMEITNDGTPTID----- 315
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
597 TNASEHRLGTGLVPAQALAAETGASSNAQDENLIEFVCLNHHSTQETTIGNFFSRAGLV 556
QY 316 ----PTEDTPGPIGSPDFGILGFGVASQNKNE 344
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
657 IITMPTGTQNTDGYVWMDIDLMGVAQMRKCE 689

RESULT 13
POLG EC01F STANDARD; PRT; 2184 AA.
AC 091734; Q66795; O9YID6;
DT 15-JUL-1999 (Rel. 38, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Genome polyprotein [Contains: Coat protein VP1 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Picornain 2A (EC 3.4.22.29) (P2A); Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P3B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase (EC 2.7.7.48) (P3D)].
OS Echovirus 1 (strain Farouk / ATCC VR-1038).
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae;
OC Enterovirus.
OX NCBI_TaxID=103908;
RN [1]_TaxID=103908;
RP SEQUENCE FROM N.A.
RA Bergelson J.M.;
RT "Receptor interactions, infectious cDNA, and nucleotide sequences of echovirus 1/8."
RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 69-330 FROM N.A.
RX MEDLINE=96203959; PubMed=8627260;
RA Huttunen P., Santti J., Pulli T., Hyypiaie T.;
RT "The major echovirus group is genetically coherent and related to coxsackie B viruses."
RL J. Gen. Virol. 77:715-725 (1996).
RN [3]
RP SEQUENCE OF 570-853 FROM N.A.
RX MEDLINE=99138973; PubMed=9971773;
```

```
RA Oberste M.S., Maher K., Kilpatrick D.R., Pallansch M.A.;
RT "Molecular evolution of the human enteroviruses: correlation of serotype with VP1 sequence and application to picornavirus classification."
RL J. Virol. 73:1941-1948 (1999).
RN [4]
RX X-RAY CRYSTALLOGRAPHY (3.55 ANGSTROMS) OF 2-850.
MEDLINE=99192672; PubMed=10089503;
RA Fillman D.J., Wien M.W., Cunningham J.A., Bergelson J.M., Hogle J.M.;
RT "Structure determination of echovirus 1."
RL Acta Crystallogr. D 54:1261-1272 (1998).
CC -!- FUNCTION: P2A AND THE P3C POLYPEPTIDES ARE PROTEASES THAT CLEAVE AT CERTAIN Q/G SITES IN THE POLYPROTEIN. THEY ARE CYSTEINE PROTEASES.
CC -!- FUNCTION: It is thought that the P2C protein attaches to vesicular membranes and is associated with viral RNA synthesis.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Tyr-Gly bond in the picornavirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + (RNA) (N).
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins. CLEAVAGE BETWEEN VP4 AND VP2 IS AUTOCATALYTIC; VP1/P2A IS CATALYZED BY P2A; ALL OTHER CLEAVAGES ARE CATALYZED BY P3C.
CC -!- SIMILARITY: P3C PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC -!- SIMILARITY: P2C PROTEASE BELONGS TO PEPTIDASE FAMILY C3.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF029859; AAC63944.2; -.
CC EMBL; X89531; CAA61710.1; -.
CC EMBL; AF081314; AAD17718.1; -.
CC PDB; 1EV1; 27-JAN-99.
CC InterPro; IPR003593; AAA_ATPase.
CC InterPro; IPR004004; Calici_pol_hel.
CC InterPro; IPR009003; Cys_Ser_trypsin.
CC InterPro; IPR000199; Pept_3C_picorn.
CC InterPro; IPR000081; Peptidase_C3.
CC InterPro; IPR003138; Pico_P1A.
CC InterPro; IPR002527; Pico_P2B.
CC InterPro; IPR001676; Rnv.
CC InterPro; IPR000605; RNA_helicase.
CC InterPro; IPR007095; RNA_pol_DS_PS.
CC InterPro; IPR001205; RNA_pol_P3D.
CC InterPro; IPR007094; RNA_pol_P5vir.
CC InterPro; IPR008975; Viral_cap_coat.
CC Pfam; PF00548; Cys-protease-3C; 1.
CC Pfam; PF02226; Pico_P1A; 1.
CC Pfam; PF00947; Pico_P2A; 1.
CC Pfam; PF01552; Pico_P2B; 1.
CC Pfam; PF00073; rhv; 3.
CC Pfam; PF00680; RNA_dep_RNA_pol; 1.
CC Pfam; PF00910; RNA_helicase; 1.
CC PRINTS; PR00918; CALICIVIRUSNS.
CC ProDom; PD001125; Cys_protease_3C; 1.
CC ProDom; PD0011306; Pico_P2A; 1.
CC ProDom; PD001274; Pico_P2B; 1.
CC SMART; SM00382; AAA; 1.
KW Polyprotein; Coat protein; Core protein; Transferase; Myristate;
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; 3D-structure;
KW Lipoprotein.
```

FT CHAIN	2	69	COAT PROTEIN VP4.	FT TURN	338	341
FT CHAIN	70	330	COAT PROTEIN VP2.	FT TURN	345	346
FT CHAIN	331	569	COAT PROTEIN VP3.	FT STRAND	353	353
FT CHAIN	570	850	COAT PROTEIN VP1.	FT TURN	356	357
FT CHAIN	851	1000	PICORNAIN 2A.	FT STRAND	369	370
FT CHAIN	1001	1099	CORE PROTEIN P2B.	FT STRAND	372	372
FT CHAIN	1100	1428	CORE PROTEIN P2C.	FT TURN	373	373
FT CHAIN	1429	1517	CORE PROTEIN P3A.	FT HELIX	374	377
FT CHAIN	1518	1539	GENOME-LINKED PROTEIN VP.	FT TURN	378	378
FT CHAIN	1540	1722	PICORNAIN 3C.	FT STRAND	381	382
FT CHAIN	1723	2184	RNA-DIRECTED RNA POLYMERASE.	FT HELIX	390	393
FT ACT SITE	1686	1686	N-myristoyl glycine (in host).	FT TURN	394	394
FT ACT SITE	1700	1700	PROTEASE 3C (POTENTIAL).	FT HELIX	397	399
FT CONFLICT	238	238	PROTEASE 3C (POTENTIAL).	FT STRAND	400	405
FT CONFLICT	611	611	N -> M (IN REF. 2).	FT TURN	406	406
FT CONFLICT	645	645	A -> R (IN REF. 2).	FT TURN	409	410
FT CONFLICT	724	724	A -> V (IN REF. 3).	FT STRAND	412	417
FT CONFLICT	774	774	Y -> F (IN REF. 3).	FT TURN	420	422
FT STRAND	4	7	V -> I (IN REF. 3).	FT HELIX	424	427
FT STRAND	26	29	A -> T (IN REF. 3).	FT TURN	428	428
FT HELIX	36	38		FT HELIX	430	435
FT STRAND	45	45		FT TURN	436	437
FT TURN	50	50		FT STRAND	438	442
FT HELIX	51	54		FT STRAND	445	452
FT STRAND	57	57				
FT TURN	63	64				
FT TURN	80	81				
FT STRAND	83	86				
FT STRAND	91	94				
FT STRAND	101	102				
FT HELIX	103	105				
FT TURN	113	115				
FT STRAND	123	123				
FT HELIX	126	128				
FT TURN	129	129				
FT STRAND	133	134				
FT STRAND	138	140				
FT TURN	142	143				
FT STRAND	147	151				
FT TURN	152	152				
FT HELIX	153	157				
FT HELIX	159	167				
FT STRAND	168	180				
FT TURN	185	186				
FT STRAND	188	197				
FT TURN	198	198				
FT STRAND	203	203				
FT TURN	206	207				
FT HELIX	212	214				
FT TURN	215	215				
FT STRAND	217	217				
FT TURN	218	219				
FT STRAND	220	220				
FT STRAND	222	223				
FT STRAND	225	225				
FT TURN	232	233				
FT STRAND	235	235				
FT STRAND	238	238				
FT TURN	239	241				
FT TURN	242	245				
FT STRAND	246	246				
FT HELIX	248	253				
FT STRAND	256	260				
FT TURN	261	263				
FT STRAND	266	271				
FT STRAND	280	280				
FT TURN	282	284				
FT STRAND	285	285				
FT STRAND	288	289				
FT TURN	302	303				
FT STRAND	308	324				

Query Match

Best Local Similarity 4.2%; Score 121.5; DB 1; Length 2184;

Mismatches 126; Conservative 73; Mismatches 178; Indels 339; Gaps 29;

QY	11	SSDGAAGLVPEINNEVM-----PLEPVAGASLATPV-----	41
Db	263	TNNSATVMPYVNSVPMNDNMYRHHNFTLMIIPFVPLDFSAGASTYVPTVTVAPMCAEYN	322
QY	42	---VGQNI-----IDPWIRNNFVQAPAG--EFTVSPR-NSPGEILLDEL-----	81
Db	323	GURLAGHOGLPMTNTPGNSQFLTSDFOPSAMPQDVTPEMHIFGEVRNLMIEIAEVS	382
QY	82	-----GPDLPN-----YLAHLARMYNGHA	100
Db	383	MPINNDAAKVSSMEAYRVELSTNINACTQVFGQLNFGAESVMNRTLMGEILNYAHWS	442
QY	101	GMGEVOIVLAGNAFTAGKIIFAAIPGPPYENLSPQITMCPHVIIDVROLEPFLPMPD	160
Db	443	GSIKITFVFCGSAMTTGKLLSYAPPAGAPK-TRKDALMLGTHVYVMDVGLQSSCVLCIP-	500
QY	161	INWNEFFHY-----NOG-----NDPKLRVAMLYTPLRANNGDDV	196
Db	501	-WISQTHRFVEKDPYTNAGFVTCWYQTSVSPASNQPKCYMCMV-----SACNDF	551
QY	197	TVSCRVLTKPSDDEF-----TFLVPPTVE-SKTQKQFALPKISEMTNSRPP	243
Db	552	SVRLMRDITKFIQTSTFYQGDVQNAVEGAMVRVADTVQTSATNSERVPNLTAVETGHTSQA	611
QY	244	VPVDVVMYTARNENQVVQPONGRVTLGDGELLGTTPLLANVICFKPGEVIAKNG---DVR	300
Db	612	VPDGTMTQTRHVINNHVRSES---TIEN-----FLARSACVFYLEY--KTGTKEDSNF	659
QY	301	-----RMDMETNTDGTPIDPTEDT-----	320
Db	660	NNWVITTRRVAQLRRKLEMFYVLRFDMEIT-----VWITSSQDQSTSQONAPVLTHQIMY	715
QY	321	-----PGPIGSPDF-----	332
Db	716	VPPGGPIFVSVDYDYSQWSTNPSIFWTEGNAPARMSIPFISIGNAYSNFYDGMHSFQAG	775
QY	333	LFGVASQRNK-----NEQNPAATRAHEAINTGGDH-----LCPOISSSEIYL	374
Db	776	VYGFITLNNMGQLFFRHVNKNPAAITSVARIYFKPKHVRVAVPPRPLCPYINSTNV--	833
QY	375	TSPNILRCTNPQPLPQSGILRGTLIRSDNGHCHDMVGTSPTTTPWQQ-----	422
Db	834	-----NFEFKPVTEVRTNII-----TTGAFQSQSGAVYVGNRVV	868

QY 423 -----WRR-----SRGSN-----CCSSGHRYPV 441  
DB 869 NRHLATHIDWQNCWEDYNDRLVSTTARGCDTIARCOCTGTVFCLSRNKHYPV 924

RESULT 14  
POLG HE71M STANDARD; PRT: 2193 AA.  
AC Q66479;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Genome polyprotein [Contains: Coat protein VP4 (P1A); Coat protein VP2 (P1B); Coat protein VP3 (P1C); Coat protein VP1 (P1D); Core protein P2A; Core protein P2B; Core protein P2C; Core protein P3A; Genome-linked protein VP3 (P2B); Picornain 3C (EC 3.4.22.28) (Protease 3C) (P3C); RNA-directed RNA polymerase P3D (EC 2.7.7.48)].  
DE Human enterovirus 71 (strain 7423/MS/87) (Ev 71).  
OS Viruses; ssRNA positive-strand viruses, no DNA stage; Picornaviridae; Enterovirus.  
OC  
OX NCBI\_TaxID=103922;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96434998; PubMed=8837884;  
RA Brown B.A., Pallansch M.A.;  
RT "Complete nucleotide sequence of enterovirus 71 is distinct from poliovirus";  
RL Virus Res. 39:195-206(1995).  
CC -!- FUNCTION: P3C POLYPEPTIDE IS A PROTEASE THAT CLEAVES AT CERTAIN Q/G SITES IN THE POLYPROTEIN. IT MAY BE A CYSTEINE PROTEASE.  
CC -!- CATALYTIC ACTIVITY: Selective cleavage of Gln-Gly bond in the poliovirus polyprotein. In other picornavirus reactions Glu may be substituted for Gln, and Ser or Thr for Gly.  
CC -!- CATALYTIC ACTIVITY: N nucleoside triphosphate = N diphosphate + {RNA} (N).  
CC -!- SUBUNIT: The virus capsid is composed of 60 icosahedral units, each of which is composed of one copy each of proteins VP1, VP2, VP3, and VP4.  
CC -!- PTM: Specific enzymatic cleavages in vivo yield mature proteins.  
CC -!- SIMILARITY: THE PROTEASE BELONGS TO PEPTIDASE FAMILY C3.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; U22522; AAB39969.1; -.  
CC HSSP; P03300; 1POV.  
DR InterPro; IPR003593; AAA\_ATPase.  
DR InterPro; IPR004004; Calci\_pol\_hel.  
DR InterPro; IPR009003; Cys\_Ser\_trypsin.  
DR InterPro; IPR000199; Pept\_3C\_picorn.  
DR InterPro; IPR000081; Peptidase\_C3.  
DR InterPro; IPR003138; Pico\_P1A.  
DR InterPro; IPR002527; Pico\_P2B.  
DR InterPro; IPR001676; Rhv.  
DR InterPro; IPR000605; RNA\_helicase.  
DR InterPro; IPR007095; RNA\_pol\_DS\_PS.  
DR InterPro; IPR001205; RNA\_pol\_P3D.  
DR InterPro; IPR007094; RNA\_pol\_Psvir.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00548; Cys-protease\_3C; 1.  
DR Pfam; PF02226; Pico\_P1A; 1.  
DR Pfam; PF00947; Pico\_P2A; 1.  
DR Pfam; PF01552; Pico\_P2B; 1.  
DR Pfam; PF00073; rhv; 3.  
DR Pfam; PF00680; RNA\_dep\_RNA\_pol; 1.  
DR Pfam; PF00910; RNA\_helicase; 1.  
DR PRINTS; PR00918; CALICVIRUSNS.  
DR ProDom; PD001125; Cys\_protease\_3C; 1.

DR ProDom; PD0011306; Pico\_P2A; 1.  
DR ProDom; PD001274; Pico\_P2B; 1.  
DR SMART; SM00382; AAA; 1.  
KW Polyprotein; Coat protein; Core protein; Core proteinase; Thiol protease; Myristate;  
KW RNA-directed RNA polymerase; Hydrolase; Thiol protease; Myristate;  
KW Lipoprotein.  
FT CHAIN 2 69 COAT PROTEIN VP4.  
FT CHAIN 70 323 COAT PROTEIN VP2.  
FT CHAIN 324 565 COAT PROTEIN VP3.  
FT CHAIN 566 862 COAT PROTEIN VP1.  
FT CHAIN 863 1012 CORE PROTEIN P2A.  
FT CHAIN 1013 1111 CORE PROTEIN P2B.  
FT CHAIN 1112 1440 CORE PROTEIN P2C.  
FT CHAIN 1441 1526 CORE PROTEIN P3A.  
FT CHAIN 1527 1548 GENOME-LINKED PROTEIN VPG.  
FT CHAIN 1549 1731 PICORNAIN 3C.  
FT CHAIN 1732 2193 RNA-DIRECTED RNA POLYMERASE P3D.  
FT LIPID 2 2 N-myristoyl glycine (in host) (By similarity).  
FT ACT\_SITE 1695 1695 PROTEASE (POTENTIAL).  
FT ACT\_SITE 1709 1709 PROTEASE (POTENTIAL).  
SQ SEQUENCE 4 2193 AA; 242656 MW; 35E1B3CF88A50EF CRC64;  
Query Match 4.1%; Score 120; DB 1; Length 2193;  
Best Local Similarity 20.5%; Pred. No. 1;  
Matches 115; Conservative 62; Mismatches 199; Indels 184; Gaps 27;  
QY 58 QAPAGEFTVSPRNSPEILLDELGLPDLNPYLALHARMYNGHAGGMEVOIVLAGNAFTAG 117  
DB 399 QAGKSELCAVFRADFGD-----GFWQSTMLGQLCGYYTQWSGSLVFTFMFTGSPMATG 452  
QY 118 KIIFAAIPGPGPYEMLSPSQIT--MCPHVIVDROLEPFLPMPDINNPFHYNQGNQDPK 175  
DB 453 KMLIAYTPPGGPG---LPKDRATAMLGTHVIMDFGLQSSVTLVIP--WISNTHY----- 500  
QY 176 LRLVAMLYTPLRANNSGDDVFTVSCRVLTKPSPDEFTFELVPTVESKTKQFALPLKIS 235  
DB 501 -----RA-HARDGVF-----DYTTGLV-----SIW 520  
QY 236 EMTNSRFPVPVDMVTARNENQVQNGRVTLDGELGTTPLLAVNICKFKGEVIKNG 295  
DB 521 YQTN--YVVPICAPNTAYILALAAQKN-----FTMKLCKDTSHIL-QTA 562  
QY 296 DVRSYRMDMEINTDGTIDP--TEDTPGIPSPDQGLFGVAGRNKQNEQNPATRAHE 353  
DB 563 SIQGRVADVIESSIGDSVSALTQALPAPTQNT-----QVSSHRLDTGEVPALQAAE 616  
QY 354 --AIINTGCDHLCPOISSSEIYLTSPNLRCTNPQPLPQSGLGRTILI-----RSDNGHC 406  
DB 617 IGASNTSDSMI-----ETRCVLNHSHTAETLTDSFFSRAGLVGEIDLPLEGTTNPNGYA 672  
QY 407 H-----DMVGTSPPTTPWQ--OWRRCSRGSNCCSSGH 437  
DB 673 NWDIDITGYAQMRKVKVELFTYMRFAEFTFVACTGTGEVVPQLQVMFVPPGA----- 725  
QY 438 RYFVPVVMNRVTWIVLSHKSGFSTRLKLPQLNLRWPLRFINPDTGRVLFEARL----- 492  
DB 726 --PKESRESLAWQTATNPSVPVKLTDPDAQVS-----VPFMSASAYQWFDGYDTFGE 778  
QY 493 HKQ-----GEITVAHTGDN-----PIVMPNGYFRFE---NWV-----NQY 526  
DB 779 HKQEKDLEYGACPNMNGTFSVRTVGSCKSKYPLVV--RIYMRMKHVRWIPRPMPENQY 836  
QY 527 -----SLAPVGTGK 535  
DB 837 LFKANPNYAGNSIKPTGTSR 856

RESULT 15  
POLG FMDV1  
ID POLG FMDV1 STANDARD; PRT: 2333 AA.  
AC P03306; Q64768; Q84750; Q84751; Q84752; Q84753; Q84754; Q84760;  
AC Q84761; Q84762; Q84763; Q84764; Q84765; Q84766; Q84767; Q84768;





GenCore version 5.1.6  
Copyright (C) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: June 1, 2004, 13:35:27 ; Search time 31.4978 Seconds  
(without alignments)  
5419.275 Million cell updates/sec

Title: US-09-926-799-11  
Perfect score: 2894  
Sequence: 1 MKMASNDAPSDGAGLVP.....VNQFSLAPVGTGKRRRVQ 541

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL.25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mmc.\*  
8: sp\_organelle.\*  
9: sp\_phage.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertibrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2193.5	75.8	540	12 Q918B0	Q918b0 norwalk-lik
2	2179.5	75.3	540	12 Q91V49	Q91v49 human calic
3	1837	63.5	537	12 Q913B7	Q913b7 human calic
4	1827	63.1	537	12 Q91H09	Q91h09 human calic
5	1816.5	62.8	536	12 Q8BAC6	Q8bac6 human calic
6	1815	62.7	537	12 Q8V765	Q8v765 norwalk vir
7	1799	62.2	537	12 Q9PY74	Q9py74 human calic
8	1743	60.2	522	12 Q9QAX2	Q9qax2 norwalk-lik
9	1688.5	58.3	548	12 Q918A1	Q918a1 norwalk-lik
10	1688.5	58.3	548	12 Q8QFK5	Q8qfk5 human calic
11	1688.5	58.3	548	12 Q8QFK3	Q8qfk3 human calic
12	1687.5	58.3	548	12 Q91722	Q91722 norwalk-lik
13	1684.5	58.2	548	12 Q8V768	Q8v768 norwalk vir
14	1679.5	58.0	548	12 Q8QFK9	Q8qfk9 human calic
15	1679.5	58.0	548	12 Q8QFK8	Q8qfk8 human calic
16	1678.5	58.0	548	12 Q8BC87	Q8bc87 human calic

17	1678.5	58.0	548	12 Q8QFL1	Q8qfl1 human calic
18	1677.5	58.0	548	12 Q916E5	Q916e5 human calic
19	1677.5	58.0	548	12 Q8QFK6	Q8qfk6 human calic
20	1676.5	57.9	548	12 Q91725	Q91725 norwalk-lik
21	1675.5	57.9	548	12 Q8QFK4	Q8qfk4 human calic
22	1674.5	57.9	548	12 Q8QFK1	Q8qfk1 human calic
23	1673.5	57.8	548	12 Q8V0P5	Q8v0p5 human calic
24	1673.5	57.8	548	12 Q8V0P4	Q8v0p4 human calic
25	1672.5	57.8	548	12 Q66296	Q66296 calicivirus
26	1671.5	57.8	548	12 Q88291	Q88291 small round
27	1671.5	57.8	548	12 Q8V771	Q8v771 norwalk vir
28	1671.5	57.8	548	12 Q8BC90	Q8bc90 human calic
29	1671.5	57.8	548	12 Q8QFK2	Q8qfk2 human calic
30	1668.5	57.7	548	12 Q91728	Q91728 norwalk-lik
31	1668.5	57.7	548	12 Q915D3	Q915d3 human calic
32	1668.5	57.7	548	12 Q915D2	Q915d2 human calic
33	1667.5	57.6	548	12 Q91V40	Q91v40 human calic
34	1665.5	57.6	548	12 Q9PYA7	Q9pya7 human calic
35	1662.5	57.4	548	12 Q8QFK7	Q8qfk7 human calic
36	1662	57.4	535	12 Q911I5	Q911i5 human calic
37	1661.5	57.4	548	12 Q918A4	Q918a4 norwalk-lik
38	1659	57.3	535	12 Q915C6	Q915c6 human calic
39	1658.5	57.3	548	12 Q8QFL2	Q8qfl2 human calic
40	1657.5	57.3	548	12 Q68291	Q68291 human calic
41	1657	57.3	535	12 Q917X7	Q917x7 norwalk-lik
42	1657	57.3	535	12 Q8V784	Q8v784 norwalk vir
43	1656	57.2	535	12 Q9QMK6	Q9qmk6 chitta viru
44	1654.5	57.2	548	12 Q8QFL0	Q8qfl0 human calic
45	1653	57.1	535	12 Q68104	Q68104 hawaii cali

## ALIGNMENTS

## RESULT 1

Q918B0 ID Q918B0 PRELIMINARY; PRT; 540 AA.  
AC Q918B0;  
DT 01-DEC-2001 (Tremblrel. 19, Created)  
DT 01-DEC-2001 (Tremblrel. 19, Last sequence update)  
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)  
DE Capsid protein.  
GN ORF2.  
OS Norwalk-like virus NLV/Gwynedd/273/1994/US.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=171841;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Gwynedd/273/1994/US;  
RX MEDLINE=97193806; PubMed=9041391;  
RA Ando T., Monroe S.S., Noel J.S., Glass R.I.;  
RT "A one-tube method of reverse transcription-PCR to efficiently amplify  
a 3-kilobase region from the RNA polymerase gene to the poly(A) tail  
of small round-structured viruses (Norwalk-like viruses).";  
RL J. Clin. Microbiol. 35:570-577(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Gwynedd/273/1994/US;  
RX MEDLINE=20266071; PubMed=10804147;  
RA Ando T., Noel J.S., Fankhauser R.L.;  
RT "Genetic classification of 'Norwalk-like viruses.'";  
RL J. Infect. Dis. 181:S336-S348(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=Hu/NLV/Gwynedd/273/1994/US;  
RA Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;  
RT Submitted (Aug-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF414400; AAL12980.1;  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_Cap\_coat.  
DR Pfam; PF00915; Calicicoat; 1.  
SQ SEQUENCE 540 AA; 59384 MW; FD10B8F00CEAFF79 CRC64;

Query Match 75.8%; Score 2193.5; DB 12; Length 540;  
 Best Local Similarity 78.5%; Pred. No. 9.5e-183; Mismatches 26; Indels 17; Gaps 4;  
 Matches 431; Conservative 26; Mismatches 75; Indels 17; Gaps 4;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVQAP 60  
 DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVQAP 60  
 QY 61 AGEFTVSPRNSGCEILLDELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120  
 DB 61 AGEFTVSPRNSGCEILLDELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120  
 QY 121 FAATPPGPPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNPFHYNQNDPKRLVA 180  
 DB 121 FAATPPGPPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNPFHYNQNDPKRLVA 180  
 QY 181 MLYTTLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPTVESKTKQFALPILKISEMTNS 240  
 DB 181 MLYTTLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPTVESKTKQFALPILKISEMTNS 240  
 QY 241 RFPVPVDMYTARNENOVQVQNGRVTLDGELLGTTPLLA VNI CKFGEVIAKNGDVRSY 300  
 DB 241 RFPVPVDMYTARNENOVQVQNGRVTLDGELLGTTPLLA VNI CKFGEVIAKNGDVRSY 300  
 QY 301 RMDMEITNDGTPIIDPTDTPGIPGSPDFQGIILFGVASORNKNEONPATRAHEAIINTGG 360  
 DB 301 RMDMEITNDGTPIIDPTDTPGIPGSPDFQGIILFGVASORNKNEONPATRAHEAIINTGG 360  
 QY 361 DHLCPQISSEIYLTSPNLRCTNPQPLPQSGLRGTI-----LIRSDNGHCHDMVGTS 413  
 DB 361 DHLCPQISSEIYLTSPNLRCTNPQPLPQSGLRGTI-----LIRSDNGHCHDMVGTS 413  
 QY 414 P-TTPTWQVQWRRRCRGSCNCCSGHRYPPVPMNRVTWIVLSHKSGFSTSKLPQLNLR 472  
 DB 414 P-TTPTWQVQWRRRCRGSCNCCSGHRYPPVPMNRVTWIVLSHKSGFSTSKLPQLNLR 472  
 QY 420 PAVAPLFPGEQILFFRSQIPSSGGHELGYMDCLVPOEWQHYQEAATAQSEV-----472  
 DB 420 PAVAPLFPGEQILFFRSQIPSSGGHELGYMDCLVPOEWQHYQEAATAQSEV-----472  
 QY 473 WPLIRFINDTGRVLFEARLHKQGITVAHTGDNPIVMPNGYFRFEAWNOFYSLAPVG 532  
 DB 473 -ALIRFINDTGRVLFEARLHKQGITVAHTGDNPIVMPNGYFRFEAWNOFYSLAPVG 531  
 QY 533 TGKGRRRVQ 541  
 DB 532 TGNRRRIQ 540

RESULT 2  
 Q91V49 PRELIMINARY; PRT; 540 AA.  
 AC Q91V49; PRELIMINARY; PRT; 540 AA.  
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Capsid protein.  
 OS Human calicivirus HU/MLV/Leeds/90/UK.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OC NCBI\_TaxID=122912;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=HU/MLV/Leeds/90/UK;  
 RX MEDLINE=20404883; PubMed=10949950;  
 RA Green J., Vinje J., Gallimore C.I., Koopmans M., Hale A.D.,  
 RA Clegg J.C., Chamberlain J., Brown D.W.G.;  
 RT "Capsid protein diversity among 'Norwalk-like' viruses.";  
 RL Virus Genes 20:227-236(2000).  
 DR EMBL; AJ277608; CAB89089.1;  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_Gap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 540 AA; 59374 MW; 5FC8FBDCBDC8DAD3 CRC64;

Query Match 75.3%; Score 2179.5; DB 12; Length 540;

Best Local Similarity 78.5%; Pred. No. 1.6e-181;  
 Matches 431; Conservative 23; Mismatches 78; Indels 17; Gaps 4;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVQAP 60  
 DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNFFVQAP 60  
 QY 61 AGEFTVSPRNSGCEILLDELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120  
 DB 61 AGEFTVSPRNSGCEILLDELGPDLNPLYLAHARMYNGHAGMEVQIVLAGNAFTAGKII 120  
 QY 121 FAATPPGPPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNPFHYNQNDPKRLVA 180  
 DB 121 FAATPPGPPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNPFHYNQNDPKRLVA 180  
 QY 181 MLYTTLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPTVESKTKQFALPILKISEMTNS 240  
 DB 181 MLYTTLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPTVESKTKQFALPILKISEMTNS 240  
 QY 241 RFPVPVDMYTARNENOVQVQNGRVTLDGELLGTTPLLA VNI CKFGEVIAKNGDVRSY 300  
 DB 241 RFPVPVDMYTARNENOVQVQNGRVTLDGELLGTTPLLA VNI CKFGEVIAKNGDVRSY 300  
 QY 301 RMDMEITNDGTPIIDPTDTPGIPGSPDFQGIILFGVASORNKNEONPATRAHEAIINTGG 360  
 DB 301 RMDMEITNDGTPIIDPTDTPGIPGSPDFQGIILFGVASORNKNEONPATRAHEAIINTGG 360  
 QY 361 DHLCPQISSEIYLTSPNLRCTNPQPLPQSGLRGTI-----LIRSDNGHCHDMVGTS 413  
 DB 361 DHLCPQISSEIYLTSPNLRCTNPQPLPQSGLRGTI-----LIRSDNGHCHDMVGTS 413  
 QY 414 P-TTPTWQVQWRRRCRGSCNCCSGHRYPPVPMNRVTWIVLSHKSGFSTSKLPQLNLR 472  
 DB 414 P-TTPTWQVQWRRRCRGSCNCCSGHRYPPVPMNRVTWIVLSHKSGFSTSKLPQLNLR 472  
 QY 420 PAVAPLFPGEQILFFRSQIPSSGGHELGYMDCLVPOEWQHYQEAATAQSEV-----472  
 DB 420 PAVAPLFPGEQILFFRSQIPSSGGHELGYMDCLVPOEWQHYQEAATAQSEV-----472  
 QY 473 WPLIRFINDTGRVLFEARLHKQGITVAHTGDNPIVMPNGYFRFEAWNOFYSLAPVG 532  
 DB 473 -ALIRFINDTGRVLFEARLHKQGITVAHTGDNPIVMPNGYFRFEAWNOFYSLAPVG 531  
 QY 533 TGKGRRRVQ 541  
 DB 532 TGNRRRRVQ 540

RESULT 3  
 Q913B7 PRELIMINARY; PRT; 537 AA.  
 AC Q913B7; PRELIMINARY; PRT; 537 AA.  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Capsid protein.  
 OS Human calicivirus NLV/Idaho Falls/378/1996/US.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OC NCBI\_TaxID=173723;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NLV/Idaho Falls/378/1996/US;  
 RX MEDLINE=99034638; PubMed=9815206;  
 RA Fankhauser R.L., Noel J.S., Monroe S.S., Ando T., Glass R.I.;  
 RT "Molecular epidemiology of 'Norwalk-like' viruses' in outbreaks of  
 RT gastroenteritis in the United States.";  
 RL J. Infect. Dis. 178:1571-1578(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NLV/Idaho Falls/378/1996/US;  
 RX MEDLINE=20266071; PubMed=10804147;  
 RA Ando T., Noel J.S., Fankhauser R.L.;  
 RT "Genetic classification of 'Norwalk-like' viruses'.";  
 RL J. Infect. Dis. 181:5336-5348(2000).  
 RN [3]  
 RP SEQUENCE FROM N.A.

RC STRAIN=NLV/Idaho Falls/378/1996/US;  
 RA Noel J.S., Fankhauser R.L., Monroe S.S., Glass R.I., Ando T.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY054299; AAL15547.1; -;  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 537 AA; 58600 MW; 70ABEACFC1F5E05D CRC64;

Query Match 63.5%; Score 1837; DB 12; Length 537;  
 Best Local Similarity 64.5%; Pred. No. 1.3e-151;  
 Matches 362; Conservative 57; Mismatches 98; Indels 44; Gaps 9;

QY 1 MKVASNDAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNNFVQAP 60  
 DB 1 MKVASNDAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNNFVQAP 60  
 QY 61 AGEFTVSPRNSPGEILLDELPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120  
 DB 61 AGEFTVSPRNSPGEILLDELPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120  
 QY 121 FAAIPGPPFYENLSQITMCHVVIDVRLQLEPFLPMPDIWNNFFHYNQNDPKRLIVA 180  
 DB 121 FAAIPGPPFYENLSQITMCHVVIDVRLQLEPFLPMPDIWNNFFHYNQNDPKRLIVA 180  
 QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPPVTSKTKQFALPILKISEMTNS 240  
 DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPPVTSKTKQFALPILKISEMTNS 240  
 QY 241 REPVPDVMYRTARNNEQVQPNQGRVTLTGELLGTTPLLVANNICKPKGEVIAK-NGDVRS 299  
 DB 241 REPVPDVMYRTARNNEQVQPNQGRVTLTGELLGTTPLLVANNICKPKGEVIAK-NGDVRS 299  
 QY 300 YRMDMEITNDGTPTDPTEDTGPISGDPFGILFGVASQRNKNQONPATRAHEAIIINTG 359  
 DB 301 YQ--LQTLNLDGSPIDTDDTGPGLGCPDFTGLLYGVASQRPD---ATRAHEAIIINTG 355  
 QY 360 GDHLCPQISSSIYLTSPNLRCTNPQPLPQSGRLGTLIRSDNGCHDMVGTSPPTPTM 419  
 DB 356 SDTAPKIGQVRFYSTSSNFER-NQTHPTPIGI-----YIEGNSDF----- 397  
 QY 420 PQWRCRSGSCCCSGHRYPV--PVMNRVTWIVLSHKSGFSTSTR-----KLPO---- 468  
 DB 398 -NQWQLPRYGGLHANNHAPLAVSPLFPGEQILFRFSFIPGASGHTNGEMDCLLPQEFVQ 456  
 QY 469 -----LNLRWPLIRFNPTDGRVLPEARLHKQGITVAHTGDNPIVMPNGYFRPEA 520  
 DB 457 HFYQEAATAQSEVALLRFVNPDTGRVLFEKSLHKQGFMTIASSGDHPIMPTNGYFRPEA 516  
 QY 521 WYNQFYSAPVGTGKRRRVQ 541  
 DB 517 WYNQFYSAPVGTGSGRRRIQ 537

RESULT 4  
 Q91H09 ID Q91H09 PRELIMINARY; PRT; 537 AA.  
 AC Q91H09;  
 DT 01-DEC-2001 (TrEMBLrel. 19, Created)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)  
 DE Major capsid protein VPI.  
 OS Human calicivirus NLV/VA97207/1997.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OX NCBI\_TaxID=165510;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=NLV/VA97207/1997;  
 RA Jiang X., Zhong W.M., Farkas T., Huang P.W., Wilton N., Barrett E.,  
 RA Fulcon D., Morrow R., Matson D.O.;  
 RT "Baculovirus expression and antigenic characterization of the capsid proteins of three Norwalk-like viruses.";

RL Submitted (JUN-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY038599; AAK84676.1; -;  
 DR InterPro; IPR004005; Calici\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.  
 SQ SEQUENCE 537 AA; 58588 MW; E6D1B1F858CA0EA7 CRC64;

Query Match 63.1%; Score 1827; DB 12; Length 537;  
 Best Local Similarity 64.2%; Pred. No. 9.8e-151;  
 Matches 360; Conservative 56; Mismatches 101; Indels 44; Gaps 9;

QY 1 MKVASNDAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNNFVQAP 60  
 DB 1 MKVASNDAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQQNIIDPWIRNNFVQAP 60  
 QY 61 AGEFTVSPRNSPGEILLDELPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120  
 DB 61 AGEFTVSPRNSPGEILLDELPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120  
 QY 121 FAAIPGPPFYENLSQITMCHVVIDVRLQLEPFLPMPDIWNNFFHYNQNDPKRLIVA 180  
 DB 121 FAAIPGPPFYENLSQITMCHVVIDVRLQLEPFLPMPDIWNNFFHYNQNDPKRLIVA 180  
 QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPPVTSKTKQFALPILKISEMTNS 240  
 DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSPDFEFTFLVPPVTSKTKQFALPILKISEMTNS 240  
 QY 241 REPVPDVMYRTARNNEQVQPNQGRVTLTGELLGTTPLLVANNICKPKGEVIAK-NGDVRS 299  
 DB 241 REPVPDVMYRTARNNEQVQPNQGRVTLTGELLGTTPLLVANNICKPKGEVIAK-NGDVRS 299  
 QY 300 YRMDMEITNDGTPTDPTEDTGPISGDPFGILFGVASQRNKNQONPATRAHEAIIINTG 359  
 DB 301 YQ--LQTLNLDGSPIDTDDTGPGLGCPDFTGLLYGVASQRPD---ATRAHEAIIINTG 355  
 QY 360 GDHLCPQISSSIYLTSPNLRCTNPQPLPQSGRLGTLIRSDNGCHDMVGTSPPTPTM 419  
 DB 356 SDTAPKIGQVRFYSTSSN-FETNQTHPTPIGI-----YIEGNSDF----- 397  
 QY 420 PQWRCRSGSCCCSGHRYPV--PVMNRVTWIVLSHKSGFSTSTR-----KLPO---- 468  
 DB 398 -NQWQLPRYGGLHANNHAPLAVSPLFPGEQILFRFSFIPGASGHTNGEMDCLLPQEFVQ 456  
 QY 469 -----LNLRWPLIRFNPTDGRVLPEARLHKQGITVAHTGDNPIVMPNGYFRPEA 520  
 DB 457 HFYQEAATAQSEVALLRFVNPDTGRVLFEKSLHKQGFMTIASSGDHPIMPTNGYFRPEA 516  
 QY 521 WYNQFYSAPVGTGKRRRVQ 541  
 DB 517 WYNQFYSAPVGTGSGRRRIQ 537

RESULT 5  
 Q98AC6 ID Q98AC6 PRELIMINARY; PRT; 536 AA.  
 AC Q98AC6;  
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)  
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
 DE Major capsid protein VPI.  
 OS Human calicivirus NLV/M7/1999/US.  
 OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
 OC Norovirus.  
 OX NCBI\_TaxID=207241;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Hu/NLV/M7/1999/US;  
 RA Vinje J.;  
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY130761; AAN05735.1; -;  
 DR InterPro; IPR004005; Calici\_coat.  
 DR InterPro; IPR008975; Viral\_Cap\_coat.  
 DR Pfam; PF00915; Calici\_coat; 1.

```
SQ SEQUENCE 536 AA; 58883 MW; C3BDC61967FC8765 CRC64;
Query Match 62.8%; Score 1816.5; DB 12; Length 536;
Best Local Similarity 64.3%; Pred. No. 8.1e-150;
Matches 367; Conservative 52; Mismatches 87; Indels 65; Gaps 12;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNNFVQAP 60
DB 1 MKMASNDATPSDDGAAGLVPEINNEVMALEPVAGASIAAPVVGQNIIDPWIRNNFVQAP 60

QY 61 AGEFTVSPRNSGELLDDLELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 AGEFTVSPRNSGELLDDLELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIL 120

QY 121 FFAIPGPPFYENLSQITMCPHVIVDVRQLEPFLPMPDIWNNFFHYNQNDPKRLVA 180
DB 121 FFAIPGPPFYENLSQAQITMCPHVIVDVRQLEPFLPMPDIRNVFYHYNQNSPKRLVA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFETFLVPPTVESKTKQFALPILKISEMTNS 240
DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFETFLVPPTVESKTKNFTLPVLVRSEMTNS 240

QY 241 RFPVVDVVMYTARNENIVQPNQGRCTTDLGELLGTTPLLAVNICKFKGEVIAKNGDVRSY 300
DB 241 RFPVVDQMYTSRNEINIVQPNQGRCTTDLGELLGTTPLQSVSICNFRGTMQAKLNEQPRY 300

QY 301 RMDMEITNDGTPIDPTEDTGPISDPFGQILFGVASORKNQNEQNPATRAHEAINTGG 360
DB 301 Q--LQTLNLDGSPIDPTDMPAPLGTDPQAMLYGVASQSRSDN--ATRAHDAQIDTAG 356

QY 361 DHLCPQI-----SSSEIYLTSPNLRCTNPOPLQPSGLRGLTILIRSDNGHCHDMVGT 413
DB 357 DTFAFKIGQVRKSSDDDFDLHP-----TKFTP-----IGNVDDQH----- 394

QY 414 PTTPTWP--QOWRRCSRSGSCSSGHRYP--VPVVMNRVTWIVLSH--KSGFSTSTR---K 465
DB 395 -----PFRQWSLPNYGGHLALNHLAPAVTLPFGCEQLFFRSHPISAGGHTDGAIDCL 448

QY 466 LQOLMLRW-----PLIRFINDTGRVLFEARLHKQGFIVVAHTGNPIVM 510
DB 449 LPQ---EWIEHFYQEAAPSQSDIALVRFINPTGRVLFEARLHKQGFIVVAHTGNPIVM 505

QY 511 PNGYFRFAAWNQFYSLAPVGTGKRRRVQ 541
DB 506 PINGYFRFAAWNPFTYLPVGTGSRRIQ 536

RESULT 6
Q8V765
ID Q8V765 PRELIMINARY; PRT; 537 AA.
AC Q8V765;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE ORF2 protein (Capsid).
OS Norwalk virus, and
OS Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=11983, 95340;
RN [1]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk virus; STRAIN=U25GII;
RA Kojima S., Kageyama T., Fukushi S., Hoshino F.B., Shinohara M.,
RA Uchida K., Natori K., Takeda N., Katayama K.;
RT "Genogroup-specific primers for detect Norwalk like virus.";
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama U25;
RA Katayama K.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
```

```
RP SEQUENCE FROM N.A.
RC SPECIES=Norwalk-like virus; STRAIN=Saitama U25;
RX MEDLINE=22192455; PubMed=12202225;
RA Katayama K., Shirato-Horikoshi H., Kojima S., Kageyama T., Oka T.,
RA Hoshino F.B., Fukushi S., Shinohara M., Uchida K., Suzuki Y.,
RA Gojobori T., Takeda N.;
RA "Phylogenetic Analysis of the Complete Genome of 18 Norwalk-like
RT Viruses.";
RL Virology 299:225-239(2002).
DR EMBL; AB067543; BAB84158.1; --
DR EMBL; AB067543; BAC11831.1; --
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 537 AA; 58969 MW; 605A68451D0308B7 CRC64;

Query Match 62.7%; Score 1815; DB 12; Length 537;
Best Local Similarity 64.0%; Pred. No. 1.1e-149;
Matches 359; Conservative 59; Mismatches 99; Indels 44; Gaps 7;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGQNIIDPWIRNNFVQAP 60
DB 1 MKMASNDAAAPSSDGAAGLVPEINHEVMAIEPVAGASLAAPVVGQNIIDPWIRNNFVQAP 60

QY 61 AGEFTVSPRNSGELLDDLELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 AGEFTVSPRNPAGREFLLDLELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKIL 120

QY 121 FFAIPGPPFYENLSQITMCPHVIVDVRQLEPFLPMPDIWNNFFHYNQNDPKRLVA 180
DB 121 FAVIPGPPFYENLSQAQITMCPHVIVDVRQLEPFLPMPDIRNTFFHYNQNSGPKRLVA 180

QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFETFLVPPTVESKTKQFALPILKISEMTNS 240
DB 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFETFLVPPTVESKTKAFTLPILKISEMTNS 240

QY 241 RFPVVDVVMYTARNENIVQPNQGRCTTDLGELLGTTPLLAVNICKFKGEVIAKNGDVRSY 300
DB 241 RFPVVDQMYTSRNEINIVQPNQGRCTTDLGELLGTTPLQPSVIGFRTGLQTLADQPNY 300

QY 301 RMDMEITNDGTPIDPTEDTGPISDPFGQILFGVASORKNQNEQNPATRAHEAINTGG 360
DB 301 TVQVHLENLDGSPVDPTDEVPAPLGTDPQALFGVISOQRSDN---ATRAHEARVNTND 357

QY 361 DHLCPQIS-----SSSEIYLTSPNLRCT-----NPQPLQPSGLRGLTILIRS 401
DB 358 PTFAPQIAQVRKPSNSDFFDNEP--IKFTPVGISVDSQNSYNQWLLPRYG----- 406

QY 402 DNGHCHDMVGTSP--TPTWPQQWRRCSRSGSCSSGHRYPVPVVMNRVTWIVLSHKSGFS 460
DB 407 --GHLTNNTHLAPSVSPWPGCEQLFFRSFWPGASGHTDGAIDCLLPQEWVAHFYQEAAT 464

QY 461 TSTRKLPQLNLRWPLIRFINDTGRVLFEARLHKQGFIVVAHTGNPIVMPNGYFRFEA 520
DB 465 AQTDV-----ALIRFVNPDTRVLFEGLHKGQFTITNSGDHPIMVMPANGYFRFEA 516

QY 521 WYNQFYSLAPVGTGKRRRVQ 541
DB 517 WYNQFYSLAPVGTGSRRIQ 537

RESULT 7
Q8PY74
ID Q8PY74 PRELIMINARY; PRT; 537 AA.
AC Q8PY74;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein (Fragment).
OS Human calicivirus strain Hu/NLW/Amsterdam/98-18/1998/NET.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=106515;
```

```

RN SEQUENCE FROM N.A.
RC STRAIN=Hu/NLIV/Amsterdam/98-18/1998/NET;
RX MEDLINE=203411555; PubMed=10878050;
RA Vinje J., Koopmans M.P.G.;
RT "Simultaneous detection and genotyping of 'Norwalk-like viruses' by
RT oligonucleotide array in a reverse line blot hybridization format.";
RL J. Clin. Microbiol. 38:2595-2601(2000).
DR EMBL; AF195848; AAF05820.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
FT NON_TER 537
SQ SEQUENCE 537 AA; 58891 MW; D8B6DD65CCA07C06 CRC64;

Query Match 62.2%; Score 1799; DB 12; Length 537;
Best Local Similarity 64.2%; Pred. No. 2.7e-148;
Matches 357; Conservative 64; Mismatches 101; Indels 34; Gaps 10;

QY 1 MKQASNDAAAPSSDGAAGLVPENNEVMPLEPVAGASLATPVVGQONIIPWIRNNFVQAP 60
Db 1 MKQASNDAAAPSSDGAAGLVPENNEVMPLEPVAGASLATPVVGQONIIPWIRNNFVQAP 60
QY 61 AGEFTVSPNSPGEILLDLELGPDLNPYLALHARMYNGHAGMEVQIVLAGNAFTAGKII 120
Db 61 AGEFTVSPNSPGEILLDLELGPDLNPYLALHARMYNGHAGMEVQIVLAGNAFTAGKII 120
QY 121 FFAIIPGFPYENLSPQITMCPHVIIIDVRLQLEPFLPMPDIWNNEFFHYNQNDPKRLVA 180
Db 121 FFAIIPGFPYENLSPQITMCPHVIIIDVRLQLEPFLPMPDIWNNEFFHYNQNDPKRLVA 180
QY 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALPILKISEMTNS 240
Db 181 MLYTPLRANNSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALPILKISEMTNS 240
QY 241 REPVVDVMTARNENOVQPNQGRVTLTGELGTTPLLAIVNICKEGEVIKNGDVRSY 300
Db 241 REPVVDVMTARNENOVQPNQGRVTLTGELGTTPLLAIVNICKEGEVIKNGDVRSY 300
QY 301 RMDMEITNDGTPIDPTDTPGIPSGDPFGILFGVASQNKNEQNPAHAEIINTGG 360
Db 301 RMDMEITNDGTPIDPTDTPGIPSGDPFGILFGVASQNKNEQNPAHAEIINTGG 360
QY 361 DHLCPQISSEIYLTSPNIRCTNPQL-----PQSLRGITLIRSDNGHCHDMV 410
Db 361 DHLCPQISSEIYLTSPNIRCTNPQL-----PQSLRGITLIRSDNGHCHDMV 410
QY 411 GTSPT-TPTWPOQ-----WRCRSGNCCSSGHRYPVPMNRVTWIVLSHKSFGSTSTRK 465
Db 411 GTSPT-TPTWPOQ-----WRCRSGNCCSSGHRYPVPMNRVTWIVLSHKSFGSTSTRK 465
QY 466 LPQLNLRLPFRINPDTGRVLFEARLHKQGITVAHTGDNPIVMPNPGYFRFEAWVNOF 525
Db 466 LPQLNLRLPFRINPDTGRVLFEARLHKQGITVAHTGDNPIVMPNPGYFRFEAWVNOF 525
QY 526 YSLAPVGTGKRRRVQ 541
Db 526 YSLAPVGTGKRRRVQ 541
QY 522 YSLAPVGTGSRRVQ 537
Db 522 YSLAPVGTGSRRVQ 537

RESULT 8
Q9QAX2 PRELIMINARY; PRT; 522 AA.
AC Q9QAX2;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein (Fragment).
DE Norwalk-like virus.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
OX NCBI_TaxID=95340;
RN [1]

```

```

RP SEQUENCE FROM N.A.
RC STRAIN=NLV114;
RX MEDLINE=20150145; PubMed=10686033;
RA Yoda T., Terano Y., Shimada A., Suzuki Y., Yamazaki K., Sakon N.,
RA Oishi I., Utagawa E.T., Okuno Y., Shibata T.;
RT "Expression of recombinant Norwalk-like virus capsid proteins using a
RT bacterial system and the development of its immunologic detection.";
RL J. Med. Virol. 60:475-481(2000).
DR EMBL; AB028246; BAA89030.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
FT NON_TER 522
SQ SEQUENCE 522 AA; 57276 MW; 814CFD6CC4DF14A8 CRC64;

Query Match 60.2%; Score 1743; DB 12; Length 522;
Best Local Similarity 63.2%; Pred. No. 2.1e-143;
Matches 345; Conservative 59; Mismatches 98; Indels 44; Gaps 7;

QY 9 APSSDGAAGLVPENNEVMPLEPVAGASLATPVVGQONIIPWIRNNFVQAPGFTVSP 68
Db 9 APSSDGAAGLVPENNEVMPLEPVAGASLATPVVGQONIIPWIRNNFVQAPGFTVSP 68
QY 69 RNSPGEILLDLELGPDLNPYLALHARMYNGHAGMEVQIVLAGNAFTAGKIIIPAAIPPGF 128
Db 69 RNSPGEILLDLELGPDLNPYLALHARMYNGHAGMEVQIVLAGNAFTAGKIIIPAAIPPGF 128
QY 129 PYENLSPQITMCPHVIIIDVRLQLEPFLPMPDIWNNEFFHYNQNDPKRLVAMLYTPLRA 188
Db 129 PYENLSPQITMCPHVIIIDVRLQLEPFLPMPDIWNNEFFHYNQNDPKRLVAMLYTPLRA 188
QY 189 NNSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALPILKISEMTNSRFPVPDV 248
Db 189 NNSGDDVFTVSCRVLTKPSDPFEFTFLVPPTVESKTKQFALPILKISEMTNSRFPVPDV 248
QY 249 MYTARNENOVQPNQGRVTLTGELGTTPLLAIVNICKEGEVIKNGDVRSYRMDMEITN 308
Db 249 MYTARNENOVQPNQGRVTLTGELGTTPLLAIVNICKEGEVIKNGDVRSYRMDMEITN 308
QY 309 TDCGTDIDTDPGIPSGDPFGILFGVASQNKNEQNPAHAEIINTGGDHLCPQIS 368
Db 309 TDCGTDIDTDPGIPSGDPFGILFGVASQNKNEQNPAHAEIINTGGDHLCPQIS 368
QY 369 -----SSEIYLTSPNIRCTNPQL-----PQSLRGITLIRSDNGHCHDM 409
Db 369 -----SSEIYLTSPNIRCTNPQL-----PQSLRGITLIRSDNGHCHDM 409
QY 410 VGTSP-TPTWPOQWRCRSGNCCSSGHRYPVPMNRVTWIVLSHKSFGSTSTRKLPQ 468
Db 410 VGTSP-TPTWPOQWRCRSGNCCSSGHRYPVPMNRVTWIVLSHKSFGSTSTRKLPQ 468
QY 469 LNLRLPFRINPDTGRVLFEARLHKQGITVAHTGDNPIVMPNPGYFRFEAWVNOF 528
Db 469 LNLRLPFRINPDTGRVLFEARLHKQGITVAHTGDNPIVMPNPGYFRFEAWVNOF 528
QY 529 APVGTG 534
Db 529 APVGTG 534
QY 517 APVGTG 522
Db 517 APVGTG 522

RESULT 9
Q918A1 PRELIMINARY; PRT; 548 AA.
AC Q918A1;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Capsid protein.
DE ORF2.
OC Norwalk-like virus NLV/New Orleans/279/1994/US.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus.
RN [1]

```

```

OX NCBI_TaxID=171844;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
RA MEDLINE=97193806; PubMed=9041391;
RX Ando T., Monroe S.S., Noel J.S., Glass R.I.;
RT "A one-tube method of reverse transcription-PCR to efficiently amplify
RT a 3-kilobase region from the RNA polymerase gene to the poly(A) tail
RT of small round-structured viruses (Norwalk-like viruses).";
RL J. Clin. Microbiol. 35:570-577 (1997).
RN [2]
RP SEQUENCE OF 100-192 FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
RX MEDLINE=98071277; PubMed=9407386;
RA Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,
RA Seto Y., Monroe S.S., Glass R.I.;
RT "Correlation of patient immune responses with genetically
RT characterized small round-structured viruses involved in outbreaks of
RT nonbacterial acute gastroenteritis in the United States, 1990 to
RT 1995.";
RL J. Med. Virol. 53:372-383 (1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Genetic classification of 'Norwalk'-like viruses.";
RL J. Infect. Dis. 181:S336-S348(2000).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=Hu/NLV/New Orleans/279/1994/US;
RX MEDLINE=20266071; PubMed=10804147;
RA Ando T., Noel J.S., Fankhauser R.L.;
RT "Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL; AF414412; AAL12989.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 5990 MW; F998C603777F8595 CRC64;

Query Match 58.3%; Score 1688.5; DB 12; Length 548;
Best Local Similarity 58.8%; Pred. No. 1.3e-138;
Matches 333; Conservative 64; Mismatches 126; Indels 43; Gaps 6;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGOQNIIDPWIRNPFVQAP 60
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGOQNIIDPWIRNPFVQAP 60
QY 61 AGEFTVSPRNSPGEIILLDLGLPDLPYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGAGGFEVQVVLGNAFTAGKVI 120
QY 121 FFAIPGPGPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNNFFHYNQNDPKRLVA 180
DB 121 FFAIPGPGPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNNFFHYNQNDPKRLVA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPPTVESKTKQFALPKISMTNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPPTVESKTKQFALPKISMTNS 240
QY 241 RFPVPVDMYTARNENOVQONGRVTLDGELGTTPLLAVNICKFGEVIAKNGDVR-- 298
DB 241 RFPVPIDLSHTSPENIVVQONGRVTLDGELGTTPLLAVNICKFGEVIAKNGDVR-- 298
QY 299 -----SYRMDMEITNTDGTPIIDTPTDTPGIGSPDFQGLFEGVASORNKNEQNAT 349
DB 299 -----SYRMDMEITNTDGTPIIDTPTDTPGIGSPDFQGLFEGVASORNKNEQNAT 349
QY 301 ADTPTPLRFLNYWHIQLDNLNGTVPYDAEDIPAPLGTDFRGKVFVGASQRNPD--TT 357
DB 301 ADTPTPLRFLNYWHIQLDNLNGTVPYDAEDIPAPLGTDFRGKVFVGASQRNPD--TT 357
QY 350 RAHEAIIINTGGDLCPQISSEIYLT-----PNILRCTNPQPLP-----QSLRG 395
DB 350 RAHEAIIINTGGDLCPQISSEIYLT-----PNILRCTNPQPLP-----QSLRG 395
QY 358 RAHEAKVDTRGRFTKLGSLVIXTDSBDFQONKPTRFTPVGVGVNDNEEFQWMLPDYS 410
DB 358 RAHEAKVDTRGRFTKLGSLVIXTDSBDFQONKPTRFTPVGVGVNDNEEFQWMLPDYS 410
QY 396 TILIRSDNGHCHDMVGTSPPTPTWPOQWRRCSCGSHRYPPVPMNVRTWVLGH 455
DB 396 TILIRSDNGHCHDMVGTSPPTPTWPOQWRRCSCGSHRYPPVPMNVRTWVLGH 455

Db 411 WSLPNYSGQFTHNNVLAPAVAPNPPGSQLFFRSQRPSSGGRSNGILDCLVPQEWVQHIFY 470
QY 456 KSGFSTSTRKLPQLNLRWPLIRFNPDTGRVLFPEARLHKQGFITVAHTGDNPIVMPNGY 515
DB 471 QESAPAQTV-----ALVRYNPDTGRVLFPEARLHKLGFMTIAKNGDSPIVPPNGY 522
QY 516 FRFAWVNVQFYSLAPVGTGKRRRVQ 541
DB 523 FRFESWVNPFTYLLAPMGTCNGRRRIQ 548

RESULT 10
Q80FKS
ID Q80FKS PRELIMINARY; PRT; 548 AA.
AC Q80FKS;
DT 01-JUN-2003 (TremBLrel. 24, Created)
DT 01-JUN-2003 (TremBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TremBLrel. 25, Last annotation update)
DE Capsid protein.
OS Human calicivirus NLV/500-01/SWE.
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;
OC Norovirus
OC NCBI_TaxID=227521;
OX [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=NLV/500-01;
RA Nilsson M., Hedlund K.-O., Thorhagen M., Johansen K., Ekspong A.,
RA Svensson L.;
RT "Evolution of human calicivirus RNA in vivo: accumulation of mutations
RT in the protruding P2 domain of the capsid leads to structural changes
RT and possibly a new phenotype.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY247438; AAP03036.1; -.
DR InterPro; IPR004005; Calici_coat.
DR InterPro; IPR008975; Viral_Cap_coat.
DR Pfam; PF00915; Calici_coat; 1.
SQ SEQUENCE 548 AA; 59856 MW; 34BC961A841B6EB8 CRC64;

Query Match 58.3%; Score 1688.5; DB 12; Length 548;
Best Local Similarity 59.4%; Pred. No. 1.3e-138;
Matches 337; Conservative 67; Mismatches 118; Indels 45; Gaps 8;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGOQNIIDPWIRNPFVQAP 60
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGOQNIIDPWIRNPFVQAP 60
QY 61 AGEFTVSPRNSPGEIILLDLGLPDLPYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120
DB 61 GGEFTVSPRNSPGEVLLNLELGPENPYLAHLARMYNGAGGFEVQVVLGNAFTAGKII 120
QY 121 FFAIPGPGPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNNFFHYNQNDPKRLVA 180
DB 121 FFAIPGPGPYENLSPQITMCPHVIIDVRQLEPFLPMPDIWNNFFHYNQNDPKRLVA 180
QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPPTVESKTKQFALPKISMTNS 240
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPPTVESKTKQFALPKISMTNS 240
QY 241 RFPVPVDMYTARNENOVQONGRVTLDGELGTTPLLAVNICKFGEVIAKNGDVR-- 296
DB 241 RFPVPIDLSHTSPENIVVQONGRVTLDGELGTTPLLAVNICKFGEVIAKNGDVR-- 296
QY 297 -----VRS--YRMDMEITNTDGTPIIDTPTDTPGIGSPDFQGLFEGVASORNKNEQNAT 349
DB 297 -----VRS--YRMDMEITNTDGTPIIDTPTDTPGIGSPDFQGLFEGVASORNKNEQNAT 349
QY 301 IDTAPRFGYHWHIQLDNLNGTVPYDAEDIPGLGTDFRGKVFVGASQRN--SDGTT 357
DB 301 IDTAPRFGYHWHIQLDNLNGTVPYDAEDIPGLGTDFRGKVFVGASQRN--SDGTT 357
QY 350 RAHEAIIINTGGDLCPQISSEIYLT-----PNILRCTNPQPLPQSGI---R 394
DB 350 RAHEAIIINTGGDLCPQISSEIYLT-----PNILRCTNPQPLPQSGI---R 394
QY 358 RAHEAKVDTRGRFTKLGSLVIXTDSBDFQONKPTRFTPVGVGVNDNEEFQWMLPDYS 417
DB 358 RAHEAKVDTRGRFTKLGSLVIXTDSBDFQONKPTRFTPVGVGVNDNEEFQWMLPDYS 417
QY 395 TILIRSDNGHCHDMVGTSPPTPTWPOQWRRCSCGSHRYPPVPMNVRTWVLGH 454
DB 418 GPI-----THNNVLAPAVAPNPPGSQLFFRSQRPSSGGRSNGILDCLVPQEWVQHIF 469
```

QY	455	HKSGFSTSTRKLPQLNLWRLPIRLINPDTGRVLFEARLHKQGITVAHTGDNPPIVMPNG	514
Db	470	YQESAPAQTQV-----ALVRVVPNDTGRVLFEAKLHKLGMWITAKDGDSPVTVPPNG	521
QY	515	YRFEEAWNQFVSLAPVCTGKGRRVQ	541
Db	522	YRFESWNPFFYTLAPMTGNGRRRIQ	548

  

RESULT 11					
Q80FK3					
ID	Q80FK3	PRELIMINARY;	PRT;	548	AA.
DT	01-JUN-2003	(TREMBLrel. 24, Created)			
DT	01-JUN-2003	(TREMBLrel. 24, Last sequence update)			
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)			
DT		Capsid protein.			
OS		Human calicivirus NLV/1312-01/SWE.			
OC		Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;			
OC		Norovirus.			
OX		NCBI_TaxID=227511;			
RP		[1]			
RP		SEQUENCE FROM N.A.			
RC		STRAIN=NLV/1312-01;			
RC		Nilsson M., Hedlund K.-O., Thorhagen M., Johansen K., Ekspong A.,			
RA		Svensson L.;			
RT		"Evolution of human calicivirus RNA in vivo: accumulation of mutations			
RT		in the protruding P2 domain of the capsid leads to structural changes			
RT		and possibly a new phenotype.";			
RL		Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.			
DR		EMBL; AY247440; AAP03038.1;			
DR		InterPro; IPR004005; Calici_coat.			
DR		InterPro; IPR008975; Viral_cap.coat.			
DR		Pfam; PF00915; Calici_coat; 1.			
SQ		SEQUENCE 548 AA; 59986 MW; 1B877FCBD0C137EC CRC64;			

Qy	455	HKSFSFSTSTRKLPQNLRLWPLIFINPDTCRVLFEARLHKQGFITVAHTGDNPIVMPENG	515
Db	470	YQESAPATQV-----ALVRYVNPDTGRVLFEAKLHKGFMPTIAKDGSPITVPPNG	521
Qy	515	YFRFRAWNFQFSLAPVGTGKRRRVQ	541
Db	522	YFRFESWNPFTIAPMGTGBGRRRIQ	548
RESULT 12			
Q91722			
ID	Q91722	PRELIMINARY; PRT; 548 AA.	
AC	Q91722;		
DT	01-DEC-2001	(TREMBLrel. 19, Created)	
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)	
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)	
DE	Capaid protein.		
GN	ORF2.		
OS	Norwalk-like virus NLV/Brattleboro/321/1995/US.		
OC	Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;		
OC	Norovirus.		
OX	NCBI_TaxID=171847;		
RN	[1]_TaxID=171847;		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Hu/NLV/Brattleboro/321/1995/US;		
RC	MEDLINE=97193806; PubMed=9041391;		
RA	Ando T., Monroe S.S., Noel J.S., Glass R.I.;		
RT	"A one-tube method of reverse transcription-PCR to efficiently amplify		
RT	a 3-kilobase region from the RNA polymerase gene to the poly(A) tail		
RT	of small round-structured viruses (Norwalk-like viruses).";		
RL	J. Clin. Microbiol. 35:570-577(1997).		
RN	[2]		
RP	SEQUENCE OF 100-192 FROM N.A.		
RC	STRAIN=Hu/NLV/Brattleboro/321/1995/US;		
RC	MEDLINE=98071277; PubMed=9407386;		
RA	Noel J.S., Ando T., Leite J.P., Green K.Y., Dingle K.E., Estes M.K.,		
RA	Seto Y., Monroe S.S., Glass R.I.;		
RT	"Correlation of patient immune responses with genetically		
RT	characterized small round-structured viruses involved in outbreaks of		
RT	nonbacterial acute gastroenteritis in the United States, 1990 to		
RT	1995.";		
RL	J. Med. Virol. 53:372-383(1997).		
RN	[3]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Hu/NLV/Brattleboro/321/1995/US;		
RC	MEDLINE=20266071; PubMed=10804147;		
RA	Ando T., Noel J.S., Fankhauser R.L.;		
RT	"Genetic classification of 'Norwalk-like viruses.'";		
RL	J. Infect. Dis. 181:S336-S348(2000).		
RN	[4]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=Hu/NLV/Brattleboro/321/1995/US;		
RC	Ando T., Seto Y., Noel J.S., Monroe S.S., Glass R.I., Fankhauser R.L.;		
RL	Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.		
DR	EMBL; AF414415; AAL12998.1; -		
DR	InterPro; IPR004005; Calici.coat.		
DR	InterPro; IPR008975; Viral_Gap_coat.		
DR	Pfam; PF00915; Calici.coat.1.		
SQ	SEQUENCE 548 AA; 59933 MW; 788AFF9AF1469158 CRC64;		
Query Match 58.3%; Score 1687.5; DB 12; Length 548;			
Best Local Similarity 59.4%; Pred. No. 1.6e-138;			
Matches 336; Conservative 63; Mismatches 124; Indels 43; Gaps 6			
Qy	1	MKMASNDAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQOQNIIDPWIRNPFVQAP	60
Db	1	MKMASNDAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVQOQNIIDPWIRNPFVQAP	60
Qy	61	AGEFTVSPRNSPGRIILLDELPGDLNPLYLAHARMYNGHAGMGVQIVLAGNAFTAGKII	120
Db	61	GGEFTVSPRNSPGVEVLNLELPGPEINPLYLAHARMYNGVAGFEVQAVLAGNAFTAGKVI	120
Qy	121	FAAIPPGPFYENLSPSQITWCPHVIDRVQLPEFLIPMPDIDMNNFFHYNQGNPKPLRLVA	180





Query Match 58.0%; Score 1679.5; DB 12; Length 548;  
Best Local Similarity 59.3%; Pred. No. 7.8e-138;  
Matches 336; Conservative 67; Mismatches 119; Indels 45; Gaps 8;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGGQNIIDPWIRNNFVQAP 60  
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGGQNIIDPWIRNNFVQAP 60

QY 61 AGEFTVSPRNSPGEILLDELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120  
DB 61 AGEFTVSPRNSPGEILLDELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120

QY 121 FAAIPPGPPYENLSQITMCPHVIVDVRQLEPFLPMPDIWNFFHYNQGNPKRLVA 180  
DB 121 FAAIPPGPPYENLSQITMCPHVIVDVRQLEPFLPMPDIWNFFHYNQGNPKRLVA 180

QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPPVETSKQFALPILKISEMTNS 240  
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPPVETSKQFALPILKISEMTNS 240

QY 241 RPPVPDVMTARNENQVQNGRVTLTGELGTTPLLANVICKFKGEV---IAKNGD- 296  
DB 241 RPPVPDVMTARNENQVQNGRVTLTGELGTTPLLANVICKFKGEV---IAKNGD- 296

QY 297 -----VRSY--RMDMEITNTDGTPIIDPTEDTPGPIGSPDFGQILFGVASQRNKNQNPAT 349  
DB 297 -----VRSY--RMDMEITNTDGTPIIDPTEDTPGPIGSPDFGQILFGVASQRNKNQNPAT 349

QY 301 IDTATPRSGFNHWHIQLDNLNGTVPDPAEDIPGLGTPDFGIVFVGASQRN---SDGTT 357  
DB 301 IDTATPRSGFNHWHIQLDNLNGTVPDPAEDIPGLGTPDFGIVFVGASQRN---SDGTT 357

QY 350 RAHEAIINTGDDHLCPOISSSEIYLTS-----PNILRCTNPQPLPQSGL---R 394  
DB 350 RAHEAIINTGDDHLCPOISSSEIYLTS-----PNILRCTNPQPLPQSGL---R 394

QY 395 GTILIRSDNGHCHDMVGTSPPTTPWQQWRRCRSGNSCCSGHRYPPVPMVNRVTWIVLS 454  
DB 395 GTILIRSDNGHCHDMVGTSPPTTPWQQWRRCRSGNSCCSGHRYPPVPMVNRVTWIVLS 454

QY 455 HKSGFSTSTRKLPQLNLRLWPLIRFINDTGRVLFEARLHKOGFITVAHTGNDPIVMPNG 514  
DB 455 HKSGFSTSTRKLPQLNLRLWPLIRFINDTGRVLFEARLHKOGFITVAHTGNDPIVMPNG 514

QY 515 YFRFEAWNOFYSLAPVGTGKRRRVQ 541  
DB 515 YFRFEAWNOFYSLAPVGTGKRRRVQ 541

QY 522 YFRFESWNPFTYLLAPMGTGNGRRRIQ 548  
DB 522 YFRFESWNPFTYLLAPMGTGNGRRRIQ 548

RESULT 15  
Q80FK8 PRELIMINARY; PRT; 548 AA.

AC Q80FK8;  
DT 01-JUN-2003 (TrEMBLrel. 24, Created)  
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Capsid protein.  
OS Human calicivirus NLV/2197-00/SWE.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Caliciviridae;  
OC Norovirus.  
OX NCBI\_TaxID=227519;  
RX [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=NLV/2197-00;  
RA Nilsson M., Hedlund K.-O., Thorhagen M., Johansen K., Ekspong A.,  
RA Svensson L.;  
RT "Evolution of human calicivirus RNA in vivo: accumulation of mutations  
in the protruding P2 domain of the capsid leads to structural changes  
and possibly a new phenotype";  
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY247435; AAP03033.1; -  
DR InterPro; IPR004005; Calici\_coat.  
DR InterPro; IPR008975; Viral\_cap\_coat.  
DR Pfam; PF00915; Calici\_coat; 1.  
SQ SEQUENCE 548 AA; 59798 MW; B904034B9F4FEC24 CRC64;

Query Match 58.0%; Score 1679.5; DB 12; Length 548;

Best Local Similarity 59.3%; Pred. No. 7.8e-138;  
Matches 336; Conservative 66; Mismatches 120; Indels 45; Gaps 8;

QY 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGGQNIIDPWIRNNFVQAP 60  
DB 1 MKMASNDAAAPSSDGAAGLVPEINNEVMPLEPVAGASLATPVVGGQNIIDPWIRNNFVQAP 60

QY 61 AGEFTVSPRNSPGEILLDELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120  
DB 61 AGEFTVSPRNSPGEILLDELGPDLNPLYLAHLARMYNGHAGMEVQIVLAGNAFTAGKII 120

QY 121 FAAIPPGPPYENLSQITMCPHVIVDVRQLEPFLPMPDIWNFFHYNQGNPKRLVA 180  
DB 121 FAAIPPGPPYENLSQITMCPHVIVDVRQLEPFLPMPDIWNFFHYNQGNPKRLVA 180

QY 181 MLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPPVETSKQFALPILKISEMTNS 240  
DB 181 MLYTPLRANSGDDVFTVSCRVLTKPSPDFEFTFLVPPVETSKQFALPILKISEMTNS 240

QY 241 RPPVPDVMTARNENQVQNGRVTLTGELGTTPLLANVICKFKGEV---IAKNGD- 296  
DB 241 RPPVPDVMTARNENQVQNGRVTLTGELGTTPLLANVICKFKGEV---IAKNGD- 296

QY 297 -----VRSY--RMDMEITNTDGTPIIDPTEDTPGPIGSPDFGQILFGVASQRNKNQNPAT 349  
DB 297 -----VRSY--RMDMEITNTDGTPIIDPTEDTPGPIGSPDFGQILFGVASQRNKNQNPAT 349

QY 301 IDTATPRSGFNHWHIQLDNLNGTVPDPAEDIPGLGTPDFGIVFVGASQRN---SDGTT 357  
DB 301 IDTATPRSGFNHWHIQLDNLNGTVPDPAEDIPGLGTPDFGIVFVGASQRN---SDGTT 357

QY 350 RAHEAIINTGDDHLCPOISSSEIYLTS-----PNILRCTNPQPLPQSGL---R 394  
DB 350 RAHEAIINTGDDHLCPOISSSEIYLTS-----PNILRCTNPQPLPQSGL---R 394

QY 395 GTILIRSDNGHCHDMVGTSPPTTPWQQWRRCRSGNSCCSGHRYPPVPMVNRVTWIVLS 454  
DB 395 GTILIRSDNGHCHDMVGTSPPTTPWQQWRRCRSGNSCCSGHRYPPVPMVNRVTWIVLS 454

QY 455 HKSGFSTSTRKLPQLNLRLWPLIRFINDTGRVLFEARLHKOGFITVAHTGNDPIVMPNG 514  
DB 455 HKSGFSTSTRKLPQLNLRLWPLIRFINDTGRVLFEARLHKOGFITVAHTGNDPIVMPNG 514

QY 515 YFRFEAWNOFYSLAPVGTGKRRRVQ 541  
DB 515 YFRFEAWNOFYSLAPVGTGKRRRVQ 541

QY 522 YFRFESWNPFTYLLAPMGTGNGRRRIQ 548  
DB 522 YFRFESWNPFTYLLAPMGTGNGRRRIQ 548

Search completed: June 1, 2004, 13:53:24  
Job time : 32.4978 secs

**THIS PAGE BLANK (USPTO)**